

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:30:13 ; Search time 71 seconds

(without alignments)  
106.976 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTKDFNLDTVSYSKDSGA.....ALMCCNKATATCHCSIHYSK 57

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298	98.0	57	11	AA05238
2	298	98.0	57	17	AA05267
3	298	98.0	57	20	AA05265
4	298	98.0	57	13	AA05208
5	291	95.7	57	13	AA05298
6	291	95.7	57	14	AA05280
7	291	95.7	57	14	AA05380
8	291	95.7	57	14	AA05312
9	291	95.7	57	20	AA05670
10	185	60.9	34	15	AA05235

11	185	60.9	34	19	AA05443
12	185	60.9	34	21	AA05174
13	185	60.9	34	23	AA05097
14	181	59.5	41	18	AA05169
15	178	58.6	34	14	AA05307
16	176	57.9	34	20	AA05435
17	157	51.6	34	11	AA05237
18	157	51.6	34	19	AA05645
19	157	51.6	34	20	AA05320
20	157	51.6	34	21	AA05174
21	128	42.1	39	18	AA05165
22	105	34.5	34	14	AA05095
23	104	34.2	34	14	AA05314
24	104	34.2	34	14	AA05372
25	104	34.2	34	15	AA05616
26	104	34.2	34	17	AA05263
27	97	31.9	34	13	AA05245
28	97	31.9	34	17	AA05275
29	97	31.9	34	23	AA05577
30	97	31.9	34	23	AA05577
31	96	31.6	34	14	AA05094
32	96	31.6	34	23	AA05776
33	94.5	31.1	34	10	AA05849
34	94.5	31.1	34	16	AA05315
35	94.5	31.1	34	16	AA05315
36	94.5	31.1	34	20	AA05430
37	93	30.6	34	21	AA05032
38	92	30.3	34	14	AA05311
39	82	27.0	34	14	AA05121
40	82	27.0	34	14	AA05017
41	82	27.0	34	16	AA05080
42	81.5	26.8	35	18	AA05160
43	72	23.7	46	23	AB02936
44	63	20.7	33	22	AB02763
45	61.5	20.2	63	22	AB02048

#### ALIGNMENTS

RESULT 1  
AA05238  
ID AA05238 standard; protein: 57 AA.  
AC AA05238;  
XX 04-AUG-1990 (first entry)  
XX Nisin precursor protein and leader peptide encoded by DNA derived from  
DE Streptococcus lactis ATCC 11454.  
XX Nisin precursor peptide; nisin leader peptide;  
KM post-translational modification; Streptococcus lactis ATCC 11454;  
XX  
OS Streptococcus lactis ATCC 11454.  
FH Key  
FT Peptide  
FT /note="leader fragment responsible for inducing post-  
FT translational modification"  
FT Protein  
FT /note="nisin precursor peptide"  
XX  
XX W0900558-A.  
XX 25-JAN-1990.  
XX  
XX 30-JUN-1989; 89MO-US02820.  
XX 05-JUL-1988; 88US-0214959.  
XX (UYMA-) UNIV OF MARYLAND.  
XX

Cationic peptide n  
Cationic peptide N  
Transplant media a  
Subtilin-nisin chl  
Lactococcus lactis  
Peptide osp-91241  
Subtilin precursor  
Cationic peptide s  
Amino acid sequenc  
Cationic peptide s  
Nisin-subtilin chl  
Nisin A/SSA. Lact  
Epidermin. Staphy  
Nisin A. Lactococ  
Sequence of the ba  
Nisin A. Lactococ  
Sequence of nisin  
Nisin 2. Lactococ  
Lactococcus lactis  
Nisin A/H270. Lac  
Lactococcus lactis  
Sequence of pre-ep  
Epil protein. Sta  
S. epidermis readi  
Amino acid sequenc  
Lactococcus lactis  
Lactococcus bacter  
Bacteriocin (Gene  
Bacteriocin LL-2.  
Hyacin M51 (lanthl  
Mutant lantibiotic  
Streptococcus poly  
Novel human diagno  
Novel human diagno

PI Hanßen NJ;  
XX  
XX WPI: 1990-051685/07.  
DR N-PSDB; AA093354.  
XX  
XX  
XX Leader peptide sequence -  
PT including post-translational modification of polypeptide(s)  
PT  
XX  
XX  
PS Disclosure; Fig 3; 19pp; English.

XX The leader peptide assists in inducing post-translational modification in  
CC  
CC a protein precursor when attached to the precursor as a leader. The  
CC precursor polypeptide contains Ser, Thr and Cys which undergo  
CC modification after translation to arrive at the mature protein ,  
CC having unusual amino acids. Tag a is claimed in the patent.

XX  
XX  
XX Sequence 57 AA;

	Query Match	Similarity	Score	DB	Length
Best Local	98.0%;	98.2%;	298;	DB 11;	57;
Matches	56;	Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

RESULT 2
AAR95267
ID AAR95267 standard; Protein: 57 AA.
XX
AC AAR95267;
XX
D7 01-AUG-1996 (first entry)
XX
DE Pre-nisin A.
XX
KW Nisin A; nisa gené; antimicrobial; preservative; antibiotic;
RK lantibiotic; protein engineering.
XX
OS Lactococcus lactis strain NIZO R5.
XX
FH Key
FT Peptide
XX
PN Location/Qualifiers
XX 1..23
XX /label= sig-peptide
XX W09616180-A1.
XX

XX	20-NOV-1995;	95MO-GB02699.
FR	19-NOV-1994;	94GB-0023404.
PA	(BIOT-) BIOTECHNOLOGY & BIOLOGICAL SCI RES COUNC.	
XX		
PI	Dodd HM, Gasson MJ;	
XX		
DR	WPI: 1996-268616/27.	
DR	N-PSDB; AAT29660.	
XX		
PT	Making cell which expresses nisin but does not contain natural nisa	
PT	gene - by providing cell with variant nisa gene, and genes for nisin	
PT	modification, secretion and immunity	
XX		
PS	Disclosure; Fig 7: 69pp; English.	
XX		
CC	The gene cluster nisaBCTIPRK (see AAT29660 and AAT29661) of Lactococcus	
CC	lactis includes the nisa gene coding for pre-nisin A (AAR95267, see	
CC	also AAR95263) and the genes for nisin modification, secretion and	
CC	immunity. nisaB (AAR95268) and nisc (AAR95270) are believed to be	
CC	involved in reactions that modify pre-nisin; nist (AAR95269) is of	
CC	similar to a transport ATPase and is involved in translocation of	

CC nisin out of the cell: nisl (AAR95271) is involved in immunity to  
CC nisin. Replacement of the natural, chromosomal copy of the nislA  
CC gene with a variant nislA gene allows prodn. of high levels of nisin  
CC A variants in *Lactococcus lactis* hosts.  
XX  
SQ Sequence 57 AA;

Query Match	98.0%	Score 298	DB 17	Length 57
Best Local Similarity	98.2%	Pred. No.	8.7e-29	
Matches 56	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

QY 1 MSTKDNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCNMKATCHCSIHVSK 57  
 |||||  
 Db 1 MSTKDNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCNMKATCHCSIHVSK 57

RESULT 3  
 AAY06665  
 ID AAY06665 standard; Protein; 57 AA.  
 XX  
 AC AAY06665;  
 XX  
 DT 09-NOV-1999 (first entry)  
 XX  
 DE Nisin A of Lactobacillus lactis.  
 XX  
 KW Nisin A; nis A; lantibiotic; animal performance; growth rate;  
 KW feed conversion; probiotic; bacteriocin.  
 OS  
 XX Lactobacillus lactis.  
 PN W09941978-A1.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PF 12-FEB-1999; 99WC-IB00250.  
 XX  
 PR 18-FEB-1998; 98GB-0003424.  
 PA (PFI2 ) PFIZER INC.  
 PA (PFI2 ) PFIZER LTD.  
 XX  
 PI Flanagan AJ, Haxell MA, Rolph TP;  
 DR WPI; 1999-527402/44.  
 DR N-PSDB; AAX87792.

Novel performance enhancing method resulting in improved growth rates and feed conversion efficiencies in animals

The present sequence represents the nisin A precursor peptide of *Lactobacillus lactis* NIZO R5. The sequence is deduced from the nls A gene (see AAX87792). Nisin A is lenthionine-containing bacteriocin. The invention relates to methods of enhancing performance in an animal by administering a bacterium capable of expressing a performance enhancing polypeptide such as nisin A or nisin Z. The enhanced performance results in improved growth rates and feed conversion efficiencies. The bacterium, which may be genetically modified to express the performance enhancing polypeptide, is administered to the gastrointestinal tract, especially to an embryo or neonatal animal. If the polypeptide is nisin, it may also inhibit ruminal methane, decrease acetate to nisin, may also inhibit ruminal methane, decrease acetate to propionate ratios and prevent amino acid deamination. Administration of an appropriate bacterium which may act as a probiotic may also help control enteric pathogens in poultry. The polypeptide is produced continuously in the gut, maintaining a constant level. The protein is eventually degraded, leaving no residues in the meat.

Sequence 57 AA;

Query Match 98.0%; Score 298; DB 20; Length 57;  
 Best Local Similarity 98.2%; Pred. No. 8.7e-29;  
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCHCSIHSK 57  
 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCHCSIHSK 57

## RESULT 4

AA03208  
 ID AA03208 standard; Protein: 57 AA.

AC AA03208;

DT 03-AUG-1999 (first entry)

DE Amino acid sequence of nisin A.

Sublancin 168; antimicrobial; food preservative; pro-sublancin 168;  
 Gram-positive bacteria; pre-sublancin 168; nisin A.

OS unknown.

PN WO9903352-A1.

PD 28-JAN-1999.

PF 17-JUL-1998; 98WO-US14547.

PR 18-JUL-1997; 97US-0053035.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Hansen JN;

DR WPI: 1999-131752/11.

PT New antimicrobial peptide, sublancin 168, from *Bacillus subtilis* -  
 used for, e.g. treatment of infections caused by Gram negative  
 bacteria and as food preservative

Dislosure: Page 53; 71pp; English.

This is the amino acid sequence of nisin A used in the method of  
 the invention involving the use of prosublancin 168. The peptide  
 designated sublancin 168, is an antimicrobial useful for treating  
 infections and preserving food against spoilage bacteria,  
 particularly Gram-positive bacteria. Pro-sublancin 168 and  
 pre-sublancin 168, are the precursors of sublancin 168. Sublancin  
 168 is very stable at low pH and can be autoclaved without damage.  
 It does not decompose after 2 years in aqueous solution of about  
 neutral pH.

Sequence 57 AA:

Query Match 98.0%; Score 298; DB 20; Length 57;  
 Best Local Similarity 98.2%; Pred. No. 8.7e-29;  
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCHCSIHSK 57  
 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCHCSIHSK 57

RESULT 5  
 ID AAR28298 standard; Protein: 57 AA.

AC AAR28298;

DT 02-APR-1993 (first entry)

DE Sequence encoded by nisz gene isolated from *L. lactis* NIZO 22186.  
 Lantibiotic; nisin Z; nisin A; analogue; food preservative.

OS Lactococcus lactis.

FT Key Location/Qualifiers  
 FT Peptide 1..23  
 FT /label= leader

PN WO9218633-A.

PD 29-OCT-1992.

PF 09-APR-1992; 92WO-NL00068.

PR 11-APR-1991; 91NL-0000634.

PA (NIZO-) NIZO STICHTING NEDERLANDS INST ZUIVELOND.

PI De Vos WM, Kuipers OP, Slezzen RJ;

DR WPI: 1992-382116/46.

DN N-PSDB: AAQ28299.

PT New lantibiotic cpds. related to nisin A - and *Lactococcus*  
 strains which produce them, useful as preservatives for foods and  
 animal feeds

Example: Fig 3; 42pp; English.

Total DNA was isolated from the *L. lactis* strains NIZO 22186 and  
 NIZO R5. With the aid of the nisa gene of strain NIZO R5 as a probe,  
 a 4.5 kb HindIII fragment was identified in the total DNA strain  
 NIZO 22186, which was then cloned in M13 mp18. The DNA sequence of  
 the gene for nisin Z production (nisZ) was determined by making use  
 of oligos complementary to the 5' and 3' flanking sequences of the  
 nisa gene. The nucleotide sequence of the nisz gene is found to be  
 identical to that of the nisa gene with the exception of a C to A  
 transversion in posn. 148 which results in the replacement of A A  
 His27 by Asn27. The above indicate that the structure of nisin Z is  
 as shown in AAR28299.

Sequence 57 AA:

Query Match 95.7%; Score 291; DB 13; Length 57;  
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;  
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCHCSIHSK 57  
 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIHSK 57

RESULT 6  
 ID AAR41280 standard; Protein: 57 AA.

AC AAR41280;

DT 10-MAR-1994 (first entry)

DE Bacteriocin.

Bacteriocin; LU-2; gram positive bacteria; antimicrobial; food;  
 inhibit.

OS Lactococcus lactis.

FT Key Location/Qualifiers  
 FT Peptide 1..23  
 FT /label= sig-peptide  
 FT Protein 24..57

```

FT      1 /label= mat_protein
FT      2 /note= "Claim 3"
FT      3 Modified-site
FT      4 /note= "The CH2 in the side chain joins
FT      5 the S in residue 30"
FT      6 Modified-site
FT      7 /note= "The S joins the side chain of residue 26"
FT      8 Modified-site
FT      9 /note= "The CH2 in the side chain joins
FT     10 the S in residue 34"
FT     11 Modified-site
FT     12 /note= "The S joins the side chain of residue 31"
FT     13 Modified-site
FT     14 /note= "The CH2 in the side chain joins
FT     15 the S in residue 42"
FT     16 Modified-site
FT     17 /note= "The S joins the side chain of residue 36"
FT     18 Modified-site
FT     19 /note= "The CH2 in the side chain joins
FT     20 the S in residue 51"
FT     21 Modified-site
FT     22 /note= "The S joins the side chain of residue 46"
FT     23 Modified-site
FT     24 /note= "The CH2 in the side chain joins
FT     25 the S in residue 52"
FT     26 Modified-site
FT     27 /note= "The S joins the side chain of residue 49"
FT     28
FT     29 US5232849-A.
FT     30 03-AUG-1993.
FT     31
FT     32 01-JUL-1991; 91US-0721774.
FT     33
FT     34 01-JUL-1991; 91US-0721774.
FT     35 PR
FT     36 14-MAY-1992; 92US-0882079.
FT     37
FT     38 (QUES-) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
FT     39
FT     40 Henderson JT, Marugg JD, Van Wasseenaar PD, Vedamuthu ER;
FT     41 WPI; 1993-287077/36.
FT     42 N-PSDB; AAQ49150.
FT     43
FT     44 Bacteriocin from Lactococcus lactis subspecies lactis - useful as
FT     45 inhibitory against Gram-positive bacteria
FT     46
FT     47 Disclosure; Page 13-14 (col 15,16,17,18); 14pp; English.
FT     48
FT     49 The sequence (AAQ49150) was amplified using primers (AAQ49151-52). The
FT     50 encoded protein inhibits selected gram positive bacteria and this
FT     51 property is enhanced if further purified by HPLC. The materials
FT     52 being treated to provide inhibition are preferably foods, although
FT     53 other materials may be treated.
FT     54
FT     55 Sequence 57 AA;
FT     56
FT     57 Query Match 95.7%; Score 291; DB 14; Length 57;
FT     58 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
FT     59 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
FT     60
FT     61 QY 1 MSTKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
FT     62 DB 1 MSTKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
FT     63
FT     64 RESULT 7
FT     65 AAR33850
FT     66 ID AAR33850 standard; Protein; 57 AA.
FT     67
FT     68 AC AAR33850;
FT     69
FT     70 DT 12-MAY-1993 (first entry)

```

```

XX      1 Bacteriocin LL-2 precursor.
XX      2
XX      3 Food treatment.
XX      4
XX      5 Lactococcus lactis sub-species lactis NRRL B-18809.
XX      6
XX      7 Key Location/Qualifiers
XX      8 Peptide 1..23
XX      9 Peptide /*note= "signal peptide"
XX     10 Peptide 24..57
XX     11 /*note= "mature peptide"
XX     12
XX     13 US5173297-A.
XX     14 22-DEC-1992.
XX     15
XX     16 01-JUL-1991; 91US-0721774.
XX     17 PR
XX     18 01-JUL-1991; 91US-0721774.
XX     19
XX     20 (QUES-) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
XX     21
XX     22 Henderson JT, Marugg JD, Van Wasseenaar PD, Vedamuthu ER;
XX     23 WPI; 1993-017533/02.
XX     24 N-PSDB; AAQ34782.
XX     25
XX     26 Inhibition of Gram-positive bacteria - using bacteriocin derived
XX     27 from Lactococcus lactis sub-species lactis NRRL B-18809
XX     28
XX     29 Disclosure; Page 13; 14pp; English.
XX     30
XX     31 The sequence is that of bacteriocin LL-2 precursor which can be
XX     32 used in a method for the inhibition of Gram-positive bacteria.
XX     33 LL-2 is especially useful for treatment of food, although other
XX     34 non-food materials may also be treated.
XX     35
XX     36 Sequence 57 AA;
XX     37
XX     38 Query Match 95.7%; Score 291; DB 14; Length 57;
XX     39 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
XX     40 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX     41
XX     42 QY 1 MSTKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
XX     43 DB 1 MSTKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
XX     44
XX     45 RESULT 8
XX     46 AAR39312
XX     47 ID AAR39312 standard; Protein; 57 AA.
XX     48
XX     49 AC AAR39312;
XX     50
XX     51 21-JAN-1994 (first entry)
XX     52
XX     53 Lactococcal bacteriocin polypeptide precursor.
XX     54
XX     55 Bacteriocin; Inhibition; polypeptide; Lactococcus lactis.
XX     56
XX     57 Lactococcus lactis (subspecies lactis).
XX     58
XX     59 Key Location/Qualifiers
XX     60 Peptide 24..57
XX     61 /*note= "Bacteriocin".
XX     62
XX     63 US5231165-A.
XX     64 27-JUL-1993.
XX     65
XX     66 01-JUL-1991; 91US-0721774.

```

PR 01-JUL-1991; 91US-0721774.  
 PR 14-MAY-1992; 92US-0882715.  
 XX  
 PA (QUES-) QUEST INT FLAVORS & FOOD INGREDIENTS CO.  
 XX  
 PI Henderson JT, Marugs JD, Vanwassenaar PD, Vedamuthu ER;  
 XX WPI: 1993-249768/31.  
 DR N-PSDB; AA046818.  
 XX  
 PT Isolated and purified polypeptide from *Lactococcus lactis* sub  
 PT species *lactis* - has inhibitory activity against gram-positive  
 PI bacteria for e.g. food etc.  
 XX  
 PS Claim 1; Column 17-18; 13pp; English.  
 XX  
 CC The isolated bacteriocin obtained from the polypeptide precursor has  
 CC an inhibitory activity against selected gram positive bacteria.  
 CC The amount of bacteriocin required to provide inhibition is 15-100  
 CC arbitrary units per gram of material. The materials being treated  
 CC with the bacteriocin to provide inhibition are especially foodstuffs.  
 SQ Sequence 57 AA;

Query Match 95.7%; Score 291; DB 14; Length 57;  
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;  
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIVSK 57  
 DB 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIVSK 57

RESULT 9  
 ID AAY06670 standard; Protein: 57 AA.  
 XX AAY06670;  
 AC AAY06670;  
 XX  
 DT 09-NOV-1999 (first entry)  
 XX  
 DE Nisin Z of *Lactobacillus lactis*.  
 XX  
 KW Nisin Z; nls Z; lantibiotic; animal performance; growth rate;  
 KW feed conversion; bacteriocin; probiotic.  
 XX  
 OS *Lactobacillus lactis*.  
 QS  
 MO9941978-A1.  
 PD 26-AUG-1999.  
 XX  
 PF 12-FEB-1999; 99WO-IB00250.  
 XX  
 PR 18-FEB-1998; 98GB-0003424.  
 XX  
 PA (PRIZ ) PRIZER INC.  
 PA (PRIZ ) PRIZER LTD.  
 PI Flanagan AJ, Haxell MA, Rolph TP;  
 XX WPI: 1999-527402/44.  
 DR N-PSDB; AAX87793.  
 XX  
 PT Novel performance enhancing method resulting in improved growth  
 PT rates and feed conversion efficiencies in animals  
 PS  
 XX Disclosure: Page 71; 79pp; English.  
 XX  
 CC This sequence represents the nisin Z peptide of *Lactobacillus*  
 CC *lactis* strain 22186. Nisin Z is a natural analogue of nisin A (see  
 CC AAY06665), a lanthionine-containing bacteriocin. The invention  
 CC relates to methods of enhancing performance in an animal by

CC administering a bacterium capable of expressing a performance  
 CC enhancing polypeptide such as nisin A or nisin Z. The enhanced  
 CC performance results in improved growth rates and feed conversion  
 CC efficiencies. The bacterium, which may be genetically modified to  
 CC express the performance enhancing polypeptide, is administered to  
 CC the gastrointestinal tract, especially to an embryo or neonatal  
 CC animal. If the polypeptide is nisin, it may also inhibit ruminal  
 CC methane, decrease acetate to propionate ratios and prevent amino  
 CC acid denaturation. Administration of an appropriate bacterium which  
 CC may act as a probiotic may also help control enteric pathogens in  
 CC poultry. The polypeptide is produced continuously in the gut,  
 CC maintaining a constant level. The protein is eventually degraded,  
 CC leaving no residues in the meat.  
 SQ Sequence 57 AA;

Query Match 95.7%; Score 291; DB 20; Length 57;  
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;  
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIVSK 57  
 DB 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIVSK 57

RESULT 10  
 ID AAR62635 standard; peptide; 34 AA.  
 XX AAR62635;  
 AC AAR62635;  
 XX  
 DT 15-JUN-1995 (first entry)  
 XX  
 DE Putative intermediate for lanthionine-contg. peptide.  
 XX  
 KW Lanthionine; methylanthionine; lantibiotic; antiviral;  
 KW immunosuppressant; antimicrobial; enzyme inhibitor.  
 XX  
 OS Synthetic.  
 XX  
 PN JP06253885-A.  
 PD 13-SEP-1994.  
 XX  
 PF 09-MAR-1993; 93JP-0048385.  
 XX  
 PR 09-MAR-1993; 93JP-0048385.  
 XX  
 PA (AJIN ) AJINOMOTO KK.  
 PA (AJIN ) AJINOMOTO KK.  
 DR WPI: 1994-329026/41.  
 XX  
 PF Prepn. of lanthionine contg. peptide(s) - useful as antimicrobial,  
 PF antiviral drugs, immunosuppressants and enzyme inhibitors  
 XX  
 PS Example 2; Page 7; 8pp; Japanese.  
 XX  
 CC This is one of 5 peptides (AAR62635-R63639) containing Cys and Ser or  
 CC Thr residues which were synthesised and tested for their usefulness  
 CC as intermediates for the preparation of peptides which include  
 CC lanthionine. Peptides 3 and 4 (AAR62637-8) produced lanthionine, while  
 CC both lanthionine and methylanthionine could be produced from peptide  
 CC 2 (AAR62636).  
 CC  
 SQ Sequence 34 AA;

Query Match 60.9%; Score 185; DB 15; Length 34;  
 Best Local Similarity 97.1%; Pred. No. 2.4e-15;  
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 24 ITTSISLCTPGCKTGALMGCMKMTATCNCISIVSK 57  
 DB 1 ITTSISLCTPGCKTGALMGCMKMTATCNCISIVSK 34

## RESULT 11

ID AAM66443 standard; peptide; 34 AA.

AC AAM66443;

DT 12-JUN-1999 (first entry)

DE Cationic peptide nisin.

KM Indolicidin analogue; resistance; cationic peptide; antibiotic;

KM bacterial infection; tolerance; antibacterial; microorganism;

KM bacteria; fungus; parasite; virus.

OS Lactococcus lactis.

PN WO9840401-A2.

SD 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA00190.

XX 25-FEB-1998; 98US-0030619.

XX 10-MAR-1997; 97US-0040649.

XX 20-AUG-1997; 97US-0915314.

XX 26-SEP-1997; 97US-0060099.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, McNicol PJ, West MHP;

XX WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing

XX activity of antibiotic or overcoming tolerance, acquired resistance

XX or inherent resistance of microorganisms

XX Disclosure; Page 10; 105pp; English.

XX AAM66393 to AAM66469 represent native cationic peptides from the

XX present invention. The present invention describes compositions and

XX methods for treating infection, especially bacterial infections. The

XX compositions and methods use cationic peptides in combination with an

XX antibiotic agent which are then administered to a patient to enhance the

XX activity of the antibiotic agent, to overcome: (a) tolerance; (b)

XX acquired resistance; and (c) inherent resistance. The combinations of

XX antibiotics and cationic peptides can provide synergistic activity

XX against a microorganism that is tolerant, inherently resistant, or has

XX acquired resistance to an antibiotic agent. They can be used for killing

XX e.g. bacteria, fungi, parasites and viruses.

XX Sequence 34 AA;

XX Query Match 60.9%; Score 185; DB 19; Length 34;

XX Best Local Similarity 97.1%; Pred. No. 2.4e-15;

XX Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 24 ITTSISLCTPGCKTGALMGCMNKTATGCHSIHVS 57

XX 1 ITTSISLCTPGCKTGALMGCMNKTATGCHSIHVS 34

XX RESULT 12

XX ID AAY91742 standard; Peptide; 34 AA.

XX AAY91742;

XX 06-JUN-2000 (first entry)

XX Cationic peptide Nisin amino acid sequence.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;

XX leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;

XX breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

XX multidrug resistance.

XX Unidentified.

XX WO9965506-A2.

XX 14-JUN-1999; 99WO-CA00552.

XX 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated

XX polyoxalkylene-modified cationic peptides, useful for treating tumours

XX Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which

XX can be used in the pharmaceutical composition of the invention. The

XX invention relates to a pharmaceutical composition containing at least one

XX activated polyoxalkylene (APO)-modified cationic peptide. The

XX modification of peptides with APO increases their activity against tumour

XX cells, including those with a multidrug resistant phenotype. The

XX pharmaceutical composition can be used to treat tumours, specifically

XX lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,

XX cervix, uterus, skin, prostate, liver and colon.

XX Sequence 34 AA;

XX Query Match 60.9%; Score 185; DB 21; Length 34;

XX Best Local Similarity 97.1%; Pred. No. 2.4e-15;

XX Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 24 ITTSISLCTPGCKTGALMGCMNKTATGCHSIHVS 57

XX 1 ITTSISLCTPGCKTGALMGCMNKTATGCHSIHVS 34

XX RESULT 13

XX ID AAU90978 standard; Peptide; 34 AA.

XX AAU90978;

XX 05-JUN-2002 (first entry)

XX Transplant media associated antimicrobial peptide #14.

XX Transplant; antimicrobial peptide; pore forming agent;

XX cell surface receptor binding compound; kidney transplant;

XX cardioplegia; organ transplant; transplant rejection.

XX Lactococcus lactis.

XX WO200209738-A1.

XX 07-FEB-2002.

XX 27-JUL-2001; 2001WO-US23785.

XX 28-JUL-2000; 2000US-221632P.

XX 17-NOV-2000; 2000US-249602P.

PR 15-MAY-2001; 2001US-290932P.  
XX  
XX (MURP/) MURPHY C J.  
XX  
XX Murphy CJ, Reid TW, Meanulty JF;  
XX  
XX WPI; 2002-268995/31.  
DR  
XX  
XX Media comprising antimicrobial polypeptides or pore forming agents  
PT and/or cell surface receptor binding compounds useful for the storage  
PT and preservation of organs prior to transplant  
XX  
XX  
PS Disclosure; Page 25; 78pp; English.  
XX  
XX The invention describes new transplant compositions comprising  
CC antimicrobial polypeptides or pore forming agents and/or cell surface  
CC receptor binding compounds. The media is capable of extending the  
CC preservation period past 72 hours and can provide organs with increased  
CC functionality upon transplant. Animals receiving kidneys stored in the  
CC media of the present invention for either three or four days had serum  
CC creatinine levels of less than half of those observed in control animals  
CC receiving kidneys stored in UW solution (defined in the specification)  
CC alone. Lower serum creatinine levels are indicative of healthier kidneys  
CC and a more preferable prognosis for the transplant patient. The media of  
CC the invention are useful for decreasing the incidence and/or severity of  
CC delayed graft function in patients receiving transplanted kidneys stored  
CC and/or treated in the media. The media may also be used in procedures  
CC such as cardioplegia. It is contemplated that transplant of healthier  
CC organs leads to a decrease in chronic rejection. This sequence represents  
CC an antimicrobial peptide studied in the development of the transplant  
CC media.  
XX  
XX  
SQ Sequence 34 AA;  
OY Query Match 60.9%; Score 185; DB 23; Length 34;  
Best Local Similarity 97.1%; Pred. No. 2.4e-15;  
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
DB 24 ITSTSLCTPGCKTGALMGCMNKTATCHCSIHVS 57  
1 ITSTSLCTPGCKTGALMGCMNKTATCHCSIHVS 34  
RESULT 14  
AAY31659  
ID AAY31659 standard; Protein; 41 AA.  
XX  
XX AAY31659;  
XX  
XX 09-NOV-1999 (first entry)  
XX  
XX Subtilin-nisin chimera.  
XX  
XX Nisin; subtilin; lantibiotic; chimera; mutant; bacteriocide;  
XX preserve/active.  
XX  
XX Chimeric - Lactococcus lactis.  
XX Chimeric - Bacillus subtilis.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..7  
FT /note= "signal peptide"  
FT Protein 8..41  
FT /note= "mature protein"  
FT Region 8..18  
FT /note= "nisin (1-11)"  
FT Region 19..41  
FT /note= "subtilin(12-32)"  
XX  
XX MO9711713-A1.  
XX  
XX 03-APR-1997.  
XX  
XX

PF 30-SEP-1996; 96MO-US15160.  
XX  
XX 28-SEP-1995; 95US-0535494.  
XX  
XX (UYMA-) UNIV MARYLAND BALTIMORE.  
XX  
XX  
XX Hansen JN;  
XX  
XX WPI; 1997-225847/20.  
DR N-PSDB; AAX87829.  
XX  
XX  
XX Lantibiotic mutants and chimera(s) - having enhanced stability and  
PT activity compared to nisin  
PT  
XX  
XX Example; Fig 2; 60pp; English.  
XX  
XX The present sequence represents a chimeric pre-peptide composed of  
CC a subtilin leader region and a subtilin-nisin fusion comprising  
CC residues 1-11 of Bacillus subtilis subtilin and residues 12-32 of  
CC Lactococcus lactis nisin. The subtilin-nisin fusion was not  
CC processed into a functional lantibiotic when expressed in B.  
CC subtilis. A heterogeneous mixture of products was produced, with  
CC none of the products having the expected properties of a correctly  
CC processed polypeptide. However, the mixture contained a minor  
CC component with a specific activity that exceeded that of nisin.  
CC The invention provides lantibiotic mutants and chimeras (see also  
CC AAY31658) having enhanced activity and stability compared to nisin  
CC and subtilin. They can be produced by cultivation of transformed  
CC host cells and used e.g. as food preservatives to treat, kill or  
CC inhibit the growth of microorganisms and/or their spores.  
XX  
XX  
SQ Sequence 41 AA;  
OY Query Match 59.5%; Score 181; DB 18; Length 41;  
Best Local Similarity 80.5%; Pred. No. 9.1e-15;  
Matches 33; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
DB 17 DSGASPRITSTSLCTPGCKTGALMGCMNKTATCHCSIHVS 57  
1 DSKITPQMKSSSPCTPGCKTGALMGCMNKTATCHCSIHVS 41  
RESULT 15  
AAR43070  
ID AAR43070 standard; peptide; 34 AA.  
XX  
XX AAR43070;  
XX  
XX 09-JUN-1994 (first entry)  
XX  
XX Lactococcus lactis bacteriocin.  
XX  
XX Lactococcus lactis bacteriocin.  
XX  
XX Lactococcus; bacteriocin; lactobacillus casei; food; antimicrobial;  
XX inhibition; contamination.  
XX  
XX Lactococcus lactis.  
XX  
XX EP573768-A.  
XX  
XX 15-DEC-1993.  
XX  
XX 28-APR-1993; 93EP-0106911.  
XX  
XX 08-MAY-1992; 92US-0880003.  
XX  
XX (UNIL ) QUEST INT BV.  
XX  
XX Henderson JT, Vandenberg PA, Vedamuthu ER;  
XX  
XX WPI; 1993-396506/50.  
XX  
XX New Lactococcus sp. for preserving foods - contains DNA encoding  
PT bacteriocin from donor L-lactis strain and recipient L lactis





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:15:30 ; Search time 15 Seconds

(without alignments)  
365.311 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTRDNLVSVSKDGA.....ALMGCKMTATCHCSIHVK 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR.73.\*

1: p1r1.\*

2: p1r2.\*

3: p1r3.\*

4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	298	98.0	57	1 NILSA
2	157	51.6	56	1 NIBSA
3	94.5	31.1	52	1 EPSED
4	91.5	30.1	52	1 EPSED
5	63	20.7	995	2 S0358
6	59.5	19.6	78	2 C82577
7	59.5	19.6	330	2 T25169
8	59.5	19.6	1574	2 T13954
9	59	19.4	456	2 S20597
10	58.5	19.2	166	2 T33370
11	58.5	19.2	772	2 D56695
12	58	19.1	565	2 T53679
13	57.5	18.9	727	2 A56879
14	57.5	18.9	771	2 S35681
15	56.5	18.6	732	2 S70355
16	56.5	18.6	456	2 A31857
17	56.5	18.6	741	2 B49555
18	56.5	18.6	1700	2 T49527
19	56	18.4	107	2 T49527
20	56	18.4	506	2 AF3300
21	56	18.4	1016	2 T29208
22	55.5	18.3	333	2 T29208
23	55.5	18.3	724	2 T47149
24	55.5	18.3	1620	2 T27283
25	55.5	18.3	1891	2 T13594
26	55.5	18.3	1920	2 T13893
27	55	18.1	101	2 F82797
28	55	18.1	267	2 C85075
29	55	18.1	369	2 S13721

30	55	18.1	374	1 A56436	alcohol dehydrogen
31	55	18.1	384	2 S25771	gsal protein - mou
32	55	18.1	551	2 S23400	bud emergence med
33	55	18.1	782	2 A10062	conserved hypotet
34	55	18.1	808	1 S62594	replication licens
35	54	17.9	124	2 T22275	hypothetical prote
36	54	17.8	172	2 T50694	transcription fact
37	54	17.8	199	2 T47716	transcription fact
38	54	17.8	343	2 T30233	ornithine cyclodea
39	54	17.8	355	2 C96651	protein T3p18.9 [1
40	54	17.8	511	2 T07787	pyruvate kinase (E
41	54	17.8	518	2 S42091	tyl(56) protein -
42	54	17.8	876	2 B82163	DNA topoisomerase
43	54	17.8	1030	2 B96568	unknown protein, 2
44	54	17.8	1046	2 B98324	probable rnf efflu
45	54	17.8	1046	2 AD2959	AcirB/AcrD/AcrF fam

## ALIGNMENTS

## RESULT 1

NILSA  
nlsin precursor - Lactococcus lactis

N:Alternate names: nlsin A; nlsin Z

C:Species: Lactococcus lactis

C>Date: 21-May-1990 #sequence\_revision 12-May-1994 #text\_change 18-Jun-1999

C:Accession: A31915; A32809; B45821; A43743; S17858; B48951; S36734; S70485; S16779;

R:Buchanan, G.W.; Banerjee, S.; Hansen, J.N.

J. Biol. Chem. 263, 16260-16266, 1988

A:Title: Structure, expression, and evolution of a gene encoding the precursor of nls

A:Reference number: A92679; MUID:89034093; PMID:3141403

A:Accession: A31915

A:Molecule type: DNA

A:Residues: 1-57 <BUC>

A:Cross-references: GB:J04057; NID:9153816; PIDN:AAA8606.1; PID:9153817

A:Experimental source: ATCC 11454

A:Note: The authors identified the species as Streptococcus lactis

R:Kalella, C.; Entian, K.D.

J. Bacteriol. 171, 1597-1601, 1989

A:Title: Nlsin, a peptide antibiotic: cloning and sequencing of the nlsA gene and pos

A:Reference number: A32809; MUID:89155467; PMID:2493449

A:Accession: A32809

A:Molecule type: DNA

A:Residues: 1-57 <KAL>

A:Cross-references: GB:M24527; NID:9341189; PIDN:AAA26948.1; PID:9530218

R:Dodd, H.M.; Horn, N.; Gasson, M.J.

J. Gen. Microbiol. 136, 535-566, 1990

A:Title: Analysis of the genetic determinant for production of the peptide antibiotic

A:Reference number: A45821; MUID:90362041; PMID:2118169

A:Accession: B45821

A:Molecule type: DNA

A:Residues: 1-57 <DOD>

A:Cross-references: GB:M2727; NID:9149440; PIDN:AAA25188.1; PID:9149441

R:Steen, M.T.; Chung, Y.J.; Hansen, J.N.

Appl. Environ. Microbiol. 57, 1181-1188, 1991

A:Title: Characterization of the nlsin gene as part of a polyclonal operon in the

A:Reference number: A43743; MUID:91282469; PMID:1905517

A:Accession: A43743

A:Molecule type: DNA

A:Residues: 1-57 <STE>

A:Cross-references: GB:M65089; NID:9149447; PIDN:AAA73038.1; PID:9149448

A:Experimental source: ATCC 11454

R:Mulders, J.W.M.; Boerigter, I.J.; Rollem, H.S.; Slezzen, R.J.; de Vos, W.M.

Eur. J. Biochem. 201, 581-584, 1991

A:Title: Identification and characterization of the lactidiotin nlsin Z, a natural nl

A:Reference number: S17858; MUID:92037612; PMID:1935953

A:Accession: S17858

A:Molecule type: DNA

A:Residues: 1-49, N', 51-57 <MUL>

A:Cross-references: EMBL:X61144; NID:944046; PIDN:CAA3440.1; PID:944047

A:Experimental source: strain NIZO 22186

A:Note: nlsin Z allelic variant; amino acid composition and structure determination b

R:Engelke, G.; Gutowski-Eckel, Z.; Hammelmann, M.; Entian, K.D.  
 Appl. Environ. Microbiol. 58, 3730-3743, 1992  
 A:Title: Biosynthesis of the lantibiotic nisin: genomic organization and membrane localization  
 A:Reference number: A48951; MUID:93128945; PMID:1482192  
 A:Accession: B48951  
 A:Molecule type: DNA  
 A:Residues: 1-57 <END>  
 A:Cross-references: GB:X68307; NID:g44040; PIDN:CAA48380.1; PID:g44042  
 A:Experimental source: strain 6F3  
 A:Title: sequence extracted from NCBI backbone (NCBI:122292, NCBI:122295)  
 R:Kulpers, O.P.; Beertshuyzen, M.M.; Slezzen, R.J.; de Vos, W.M.  
 Eur. J. Biochem. 216, 281-291, 1993  
 A:Title: Characterization of the nisin gene cluster nlsABPCPR of *Lactococcus lactis*.  
 A:Reference number: S36734; MUID:9373937; PMID:7689965  
 A:Accession: S36734  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-57 <KUN>  
 A:Cross-references: GB:L16226; NID:g400365; PIDN:AAA25189.1; PID:g400366  
 R:Gross, E.; Morell, J.L.  
 Am. Chem. Soc. 93, 4634-4635, 1971  
 A:Title: The structure of nisin.  
 A:Reference number: A54460; MUID:72072901; PMID:5131162  
 A:Contents: annotation  
 R:Kulpers, O.P.; Rollema, H.S.; de Vos, W.M.; Slezzen, R.J.  
 FEBS Lett. 330, 23-27, 1993  
 A:Title: Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*, dir  
 A:Reference number: S36142; MUID:93380562; PMID:8370453  
 A:Contents: annotation  
 R:van der Meer, J.R.; Polman, J.; Beertshuyzen, M.M.; Slezzen, R.J.; Kulpers, O.P.; De Vos  
 J. Bacteriol. 175, 2578-2588, 1993  
 A:Title: Characterization of the *Lactococcus lactis* nisin A operon genes nisp, encoding  
 involved in nisin biosynthesis.  
 A:Reference number: A40621; MUID:93239683; PMID:8478324  
 A:Contents: annotation  
 R:Chan, W.C.; Leyland, M.; Clark, J.; Dodd, H.M.; Lian, L.Y.; Yang, J.C.; Roberts, G.C.K.  
 FEBS Lett. 390, 129-132, 1996  
 A:Title: Structure-activity relationships in the peptide antibiotic nisin: antibacterial  
 A:Reference number: S70485; MUID:96305786; PMID:8706842  
 A:Accession: S70485  
 A:Molecule type: protein  
 A:Residues: 24-57 <CHN>  
 C:Comment: Nisin is secreted as an inactive precursor and then activated by cleavage with  
 C:Genetics:  
 A:Gene: span; nisa  
 C:Superfamily: subtilin precursor  
 C:Keywords: antibiotic; lantibiotic  
 F:2-23/Domain: propeptide #status experimental <PRO>  
 C:Keywords: antibiotic; lantibiotic  
 F:2-24/Cleavage site: Arg-116 (protease nisp) #status experimental  
 F:2-25/Modified site: dehydrobutyryne (Thr) #status experimental  
 F:26-30/Cross-link: sn-(25,6R)-lanthionine (Ser-Cys) #status experimental  
 F:31-34/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:35-42/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:46-49/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:48-51/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:56/Modified site: dehydroalanine (Ser) #status experimental

Query Match 98.0%; Score 298; DB 1; Length 57;  
 Best local Similarity 98.2%; Pred. No. 6; 3e-27;  
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTRDENLDVSVKSKDSCASPRITSTSLCTPGCKTGALMGCMKATATGCHSIHSK 57  
 |||||||  
 DB 1 MSTRDENLDVSVKSKDSCASPRITSTSLCTPGCKTGALMGCMKATATGCHSIHSK 57  
 |||||||

RESULT 2  
 subtilin precursor - *Bacillus subtilis* (strain ATCC 6633)  
 N:Alternate names: spasi; subtilin A  
 N:Contains: subtilin B

C:Species: *Bacillus subtilis*  
 C:Date: 30-Sep-1989 #sequence-revision 12-May-1994 #text-change 21-Jul-2000  
 C:Accession: A28112; D42655; D43935; I40514; I39980; S36142  
 R:Banerjee, S.; Hansen, J.N.  
 J. Biol. Chem. 263, 9508-9514, 1988  
 A:Title: Structure and expression of a gene encoding the precursor of subtilin, a sma  
 A:Reference number: A28112; MUID:88243844; PMID:2837490  
 A:Accession: A28112  
 A:Molecule type: DNA  
 A:Residues: 1-56 <BNM>  
 A:Cross-references: GB:J03767; NID:g143718; PIDN:AAA22841.1; PID:g143719  
 R:Chung, Y.J.; Steen, M.T.; Hansen, J.N.  
 J. Bacteriol. 174, 1417-1422, 1992  
 A:Title: The subtilin gene of *Bacillus subtilis* ATCC 6633 is encoded in an operon tha  
 A:Reference number: A42655; MUID:92138640; PMID:1735728  
 A:Accession: D42655  
 A:Molecule type: DNA  
 A:Residues: 1-56 <CHN>  
 A:Cross-references: GB:M83944; NID:g143557; PIDN:AAA22772.1; PID:g143561  
 A:Experimental source: ATCC 6633  
 A:Note: sequence extracted from NCBI backbone (NCBI:79670)  
 R:Klein, C.; Kaletta, C.; Schnell, N.; Entian, K.D.  
 Appl. Environ. Microbiol. 58, 132-142, 1992  
 A:Title: Analysis of genes involved in biosynthesis of the lantibiotic subtilin.  
 A:Reference number: A43935; MUID:92171481; PMID:1539969  
 A:Accession: D43935  
 A:Molecule type: DNA  
 A:Residues: 1-56 <KLE1>  
 A:Cross-references: GB:M86869; NID:g143713; PIDN:AAA22840.1; PID:g143717  
 A:Experimental source: ATCC 6633  
 A:Note: sequence extracted from NCBI backbone (NCBI:84011, NCBI:84020)  
 R:Klein, C.; Entian, K.D.  
 Appl. Environ. Microbiol. 60, 2793-2801, 1994  
 A:Title: Genes involved in self-protection against the lantibiotic subtilin produced  
 A:Reference number: I40511; MUID:94368094; PMID:8085823  
 A:Accession: I40511  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-56 <KLE2>  
 A:Cross-references: EMBL:U09819; NID:g2702240; PIDN:AB91589.1; PID:g595319  
 A:Experimental source: ATCC 6633  
 R:Chan, W.C.; Bycroft, B.W.; Leyland, M.L.; Lian, L.Y.; Yang, J.C.; Roberts, G.C.K.  
 FEBS Lett. 300, 56-62, 1992  
 A:Title: Sequence-specific resonance assignment and conformational analysis of subtil  
 A:Reference number: A44571; MUID:92192284; PMID:1547888  
 A:Contents: annotation  
 R:Chan, W.C.; Bycroft, B.W.; Leyland, M.L.; Lian, L.Y.; Roberts, G.C.K.  
 Biochem. J. 291, 23-27, 1993  
 A:Title: A novel post-translational modification of the peptide antibiotic subtilin:  
 A:Reference number: A53265; MUID:93228611; PMID:8471040  
 R:Kulpers, O.P.; Rollema, H.S.; de Vos, W.M.; Slezzen, R.J.  
 FEBS Lett. 330, 23-27, 1993  
 A:Title: Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*,  
 A:Reference number: S36142; MUID:93380562; PMID:8370453  
 A:Contents: annotation  
 C:Comment: *Bacillus subtilis* strain ATCC 6633 carries this gene that is not found in  
 C:Genetics:  
 A:Gene: spasi  
 C:Superfamily: subtilin precursor  
 C:Keywords: antibiotic; blocked amino end; lantibiotic  
 F:2-24/Domain: propeptide #status predicted <SIG>  
 F:2-25/Modified site: dehydrobutyryne (Thr) #status experimental  
 F:25-56/Domain: subtilin B #status experimental <MATB>  
 F:25-56/Product: subtilin B #status experimental  
 F:27-31/Cross-link: sn-(25,6R)-lanthionine (Ser-Cys) #status experimental  
 F:32-35/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:36-43/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:42/Modified site: (2)-dehydrobutyryne (Thr) #status experimental  
 F:47-50/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:49-52/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental  
 F:55/Modified site: dehydroalanine (Ser) #status experimental

C;Accession: C82577  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence



```

2Y      39 LM----GCNMKTATCHCSIH 54
          |      |      |      |      |
Ddb     486 LVDPKRGTKPELIVCNCTMH 505

```

RESULT 15  
S70355

RESULT 15  
57035  
phosphoribosylanthranilate isomerase (EC 5.3.1.24) - *Lipomyces starkeyi*  
C:Species: *Lipomyces starkeyi*  
C:Date: 06-Dec-1996 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S70355  
R:Biggell, G.R.; Bruce, I.J.; Evans, I.H.  
Curr. Genet. 30, 83-88, 1996  
A:Title: Electrophoretic karyotype of the amyloidic yeast *Lipomyces starkeyi* and clonin  
A:Reference number: S70355; MUID:96269934; PMID:8662214  
A:Accession: S70355  
A:Molecule type: DNA  
A:Residues: 1-232 <BIG>  
A:Cross-references: EMBL:Z68292; NID:g1134847; PIDN:CAA92584.1; PID:g1161576  
A:Note: the authors translated the codon GCC for residue 211 as Glu  
C:Genetics:  
A:Gene: TRP1  
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology  
C:Keywords: Intramolecular oxidoreductase; isomerase; tryptophan biosynthesis  
F:14-228/Domain: trpF homology <TRF>

Query Match 18.6%; Score 56.5; DB 2; Length 232;  
Best local Similarity 36.7%; Pred. No. 41;  
Matches 11: Conservative 5; Mismatches 9; Indels 5; Gaps 1;

24 ITSTSLCTPGCKTGALMGCMKMTATCHCSI 53  
:::||||| | : |||  
3 VSTTSLCTPIVKI-----CGLTVEAHCAL 27

Search completed: June 7, 2003, 15:19:26  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:12:29 ; Search time 11 Seconds

(without alignments)  
214,923 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304 1 MSTKDFNLDLVSVSKDSCA.....ALMGCMKATCCHSYHVK 57

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwisProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298	98.0	57	1	PI1068 lactococcus
2	291	95.7	57	1	LANN_LACLA
3	157	51.6	56	1	P23559 lactococcus
4	94.5	31.1	52	1	PI0946 bacillus su
5	91.5	30.1	52	1	P08136 staphylococ
6	72	23.7	46	1	P21838 staphylococ
7	63	20.7	46	1	SRV41 streptococ
8	61	20.1	45	1	Q94421 streptococ
9	59	19.4	45	1	P40442 saccharomy
10	58.5	19.2	64	1	O9hcm4 homo sapien
11	58.5	19.2	722	1	P29315 rattus norv
12	57.5	18.9	727	1	P55947 helix pomat
13	57.5	18.9	727	1	004726 homo sapien
14	57.5	18.9	771	1	P51556 rattus norv
15	56.5	18.6	1069	1	P08122 mus musculu
16	56.5	18.6	232	1	P97435 mus musculu
17	56.5	18.6	456	1	001128 lipomyces s
18	56.5	18.6	498	1	PI0775 sus scrofa
19	56.5	18.6	741	1	006846 mus musculu
20	56.5	18.6	766	1	007141 rattus norv
21	56.5	18.6	799	1	062441 mus musculu
22	56.5	18.6	1700	1	027591 drosophila
23	56.5	18.6	3674	1	003376 chironomus
24	56	18.4	63	1	09arc6 homo sapien
25	56	18.4	71	1	066586 streptococ
26	55	18.3	3110	1	P58917 conus caten
27	55	18.3	766	1	P24043 homo sapien
28	55	18.1	369	1	004727 homo sapien
29	55	18.1	374	1	P21551 ambystoma m
30	55	18.1	384	1	064437 mus musculu
31	54.5	17.9	551	1	001721 mus musculu
32	54.5	17.9	64	1	P29366 saccharomy
33	34	17.8	166	1	P55955 streptococ
					066478 human enter
					P97615 rattus norv

## ALIGNMENTS

RESULT 1	ID	LANN_LACLA	STANDARD:	PRT:	57 AA.	
AC	PI3068;					
DT	01-JAN-1990 (Rel. 13, Created)					
DT	01-JAN-1990 (Rel. 13, Last sequence update)					
DT	16-OCT-2001 (Rel. 40, Last annotation update)					
DE	Lantibiotic nisin A precursor.					
GN	SPAN OR NISA.					
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).					
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.					
OX	NCBI_TaxID=1360;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-ATCC 11454 / DSM 20729 / NCD0 496;					
RX	MEDLINE-89034093; PubMed-3141403;					
RA	Buchanan G.W., Banerjee S., Hansen J.N.;					
RT	"Structure, expression, and evolution of a gene encoding the					
RL	precursor of nisin, a small protein antibiotic.";					
RL	J. Biol. Chem. 263:16260-16266(1988).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-ATCC 11454 / DSM 20729 / NCD0 496;					
RX	MEDLINE-91282469; PubMed-1905517;					
RA	Steen M.T., Chung Y.J., Hansen J.N.;					
RT	"Characterization of the nisin gene as part of a polycistronic operon					
RL	in the chromosome of Lactococcus lactis ATCC 11454.";					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-6F3;					
RX	MEDLINE-89155467; PubMed-2493449;					
RA	Kaletta C., Entian K.-D.;					
RT	"Nisin, a peptide antibiotic: cloning and sequencing of the nisa gene					
RL	and posttranslational processing of its peptide product.";					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-6F3;					
RX	MEDLINE-93128945; PubMed-1482192;					
RA	Engelke G., Gutowski-Eckel Z., Hammelmann M., Entian K.-D.;					
RT	"Biosynthesis of the lantibiotic nisin: genomic organization and					
RL	membrane localization of the nisa protein.";					
RN	[5]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-NIZO R5;					
RX	MEDLINE-93373937; PubMed-7689965;					
RA	Kuipers O.P., Beerthuyzen M.M., Steen R.J., de Vos W.M.;					
RT	"Characterization of the nisin gene cluster nisaATCPR of Lactococcus					
RL	lactis. Requirement of expression of the nisa and nisi genes for					
RN	development of immunity.";					
RL	Eur. J. Biochem. 216:281-291(1993).					
RN	[6]					
RP	SEQUENCE OF 24-57.					
RA	Gross E.;					

042806 glycine max  
P25205 homo sapien  
O9krb2 vibrio chol  
P29374 homo sapien  
065900 c genome po  
099757 homo sapien  
084713 chlamydia t  
Q21313 caenorhabd1  
P40529 saccharomy  
P22816 drosophila  
P41682 rattus norv  
P43220 homo sapien

RL (in) Friedman M. (eds.):  
 RL Protein cross-linking, pp.131-153, Plenum Press, New York (1977).  
 RL (7)  
 RP SEQUENCE OF 24-57.  
 RX MEDLINE-72072901; PubMed-5111162;  
 RA Gross E., Morell J.L.;  
 RT "The structure of nisin."  
 RL J. Am. Chem. Soc. 93:4634-4635(1971).  
 RN (8)  
 RP STRUCTURE BY NMR.  
 RX MEDLINE-92111494; PubMed-1765078;  
 RA van de Ven F.J., van den Hooven H.W., Konings R.N.H., Hilbers C.W.;  
 RT "NMR studies of lantibiotics. The structure of nisin in aqueous  
 RT solution."  
 RL Eur. J. Biochem. 202:1181-1188(1991).  
 RN (9)  
 RP STRUCTURE BY NMR.  
 RX MEDLINE-9246867; PubMed-1575686;  
 RA Llan L.-Y., Chan W.C., Morley S.D., Roberts G.C.K., Bycroft B.W.,  
 RA Jackson D.;  
 RT "Solution structures of nisin A and its two major degradation  
 RT products determined by NMR."  
 RL Biochem. J. 283:413-420(1992).  
 RN (10)  
 RP STRUCTURE BY NMR.  
 RX MEDLINE-93202265; PubMed-8454055;  
 RA van den Hooven H.W., Fogolari F., Rolletta H.S., Konings R.N.H.,  
 RA Hilbers C.W., van de Ven F.J.;  
 RT "NMR and circular dichroism studies of the lantibiotic nisin in non-  
 RT aqueous environments."  
 RL FEBS Lett. 319:189-194(1993).  
 RN (11)  
 RP STRUCTURE BY NMR.  
 RX MEDLINE-93120109; PubMed-8418850;  
 RA Saller W., Helms G.L., Henkel T., Niemczura W.P., Stiles M.E.,  
 RA Veders J.C.;  
 RT "15N- and 13C-labeled media from *Anabaena* sp. for universal isotopic  
 RT labeling of bacterioclams: NMR resonance assignments of leucocin A  
 RT from leucocostoc gelidium and nisin A from *Lactococcus lactis*."  
 RL Biochemistry 32:310-318(1993).  
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)  
 CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF  
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL  
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS  
 CC TRANSMEMBRANE PORES.  
 CC -1- PM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF  
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE  
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND  
 CC CLEAVAGE OF THE MODIFIED PRECURSOR.  
 CC -1- MISCELLANEOUS: USED AS A FOOD PRESERVATIVE.  
 CC -1- MISCELLANEOUS: THE NISA GENE IS FOUND BOTH ON CHROMOSOMAL AND  
 CC PLASMID DNA. THE SEQUENCES REPORTED ARE ABSOLUTELY IDENTICAL.  
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J04057; AAA88606.1;  
 DR EMBL: M65089; AAA73038.1;  
 DR EMBL: M24527; AAA26948.1;  
 DR EMBL: X68307; CAA48380.1;  
 DR EMBL: M27277; AAA25188.1;  
 DR EMBL: D00696; BAA00602.1;  
 DR EMBL: L16226; AAA25189.1;  
 DR EMBL: M79445; AAA25198.1;  
 DR PIR: A1915; NIIASA.  
 DR PIR: B48951; B48951.  
 DR PIR: S36734; S36734.

DR InterPro: IPR001049; Gallidermin.  
 DR InterPro: IPR000446; Nisin.  
 DR Pfam: PF02052; Gallidermin; 1.  
 DR PRINTS: PR00324; NISIN.  
 KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.  
 FT PROPEP 1 23  
 FT CHAIN 24 57  
 FT MOD.RES 25 25  
 FT MOD.RES 26 26  
 FT MOD.RES 28 28  
 FT MOD.RES 31 31  
 FT MOD.RES 36 36  
 FT MOD.RES 46 46  
 FT MOD.RES 48 48  
 FT MOD.RES 56 56  
 FT THIOETH 31 34  
 FT THIOETH 36 42  
 FT THIOETH 46 49  
 FT THIOETH 48 51  
 SO SEQUENCE 57 AA; 5963 MW; 315E4428AC70BFBA CRC64;  
 Query Match 98.0%; Score 298; DB 1; Length 57;  
 Best Local Similarity 98.2%; Pred No. 2e-30; 1; Indels 0; Gaps 0;  
 Matches 56; Conservative 0; Mismatches 1;  
 Oy 1 MSTKDFNLDVSVSKKSGASPRITSTSLCTPGCKTGALMGCMKTPATCCHSIHVK 57  
 Db 1 MSTKDFNLDVSVSKKSGASPRITSTSLCTPGCKTGALMGCMKTPATCCHSIHVK 57  
 RESULT 2  
 LANTZ\_LACIA STANDARD; PRT; 57 AA.  
 AC P29559;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Lantibiotic nisin Z precursor.  
 GN NISZ.  
 OS *Lactococcus lactis* (subsp. *lactis*) (*Streptococcus lactis*).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; *Lactococcus*.  
 OX NCBI\_TaxID=1360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIZO 22186;  
 RX MEDLINE-92037612; PubMed-1935953;  
 RA Mulders J.W.M., Boerrigter I.J., Rolletta H.S., Slezan R.J.,  
 RA de Vos W.M.;  
 RT "Identification and characterization of the lantibiotic nisin Z, a  
 RT natural nisin variant."  
 RL Eur. J. Biochem. 201:581-584(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JCM 7638;  
 RA Araya T., Ishibashi N., Shimamura S.;  
 RT "Genetic evidence that *Lactococcus lactis* JCM7638 produces a mutated  
 RT form of nisin."  
 RL J. Gen. Appl. Microbiol. 38:271-278(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N8;  
 RX MEDLINE-95352820; PubMed-7626780;  
 RA Immonen T., Ye S., Ra R., Qiao M., Paulin L., Sarris P.E.J.;  
 RT "The codon usage of the *nisz* operon in *Lactococcus lactis* N8 suggests  
 RT a non-lactococcal origin of the conjugative nisin-sucrose  
 RT transposon."  
 RL DNA Seq. 5:203-218(1995).  
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)  
 CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF  
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL  
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS  
 CC TRANSMEMBRANE PORES.



CC -1- P.T.M.: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF  
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE  
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND  
 CC CLEAVAGE OF THE MODIFIED PRECURSOR.  
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X61144; CAA43440.1; -  
 CC EMBL: D10768; BAA01598.1; -  
 CC EMBL: 218947; CAA79467.1; -  
 CC PIR: A31915; NILSA.  
 CC InterPro: IPR001049; Gallidermln.  
 CC InterPro: IPR000446; Nisin.  
 CC Pfam: PF02052; Gallidermln; 1.  
 CC PRINTS: PR00324; Nisin.  
 CC Antibiotic; Bacteriocin; Lantibiotic.  
 CC CHAIN 1 23  
 CC FT PROPEP 24 57  
 CC FT MOD\_RES 25 57 LANTIBIOTIC NISIN 2.  
 CC FT MOD\_RES 26 57 DHB (2,3-DIDEHYDROBUTYRINE).  
 CC FT MOD\_RES 28 57 D-ALANINE.  
 CC FT MOD\_RES 31 31 DHA (2,3-DIDEHYDROALANINE).  
 CC FT MOD\_RES 36 36 D-ABU (AMINO BUTYRIC ACID).  
 CC FT MOD\_RES 46 36 D-ABU (AMINO BUTYRIC ACID).  
 CC FT MOD\_RES 48 48 D-ABU (AMINO BUTYRIC ACID).  
 CC FT MOD\_RES 56 56 DHA (2,3-DIDEHYDROALANINE).  
 CC FT THIOETH 31 34 ALA-S-CYS (LANTHIONINE).  
 CC FT THIOETH 36 42 ABU-S-CYS (BETA-METHYLLANTHIONINE).  
 CC FT THIOETH 46 49 ABU-S-CYS (BETA-METHYLLANTHIONINE).  
 CC FT THIOETH 48 51 ABU-S-CYS (BETA-METHYLLANTHIONINE).  
 CC FT THIOETH 50 50 N -> H (IN STRAIN JCM7638).  
 CC FT VARIANT 50 50 DF5E4428AC70BEEF CRC64;  
 CC SEQUENCE 57 AA: 5940 MW: 125442.84  
 CC Query Match 95.78; Score 291; DB 1; Length 57;  
 CC Best Local Similarity 96.58; Pred. No. 1.5e-29;  
 CC Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC 1 MSTRDPNLDVSVSKDGSAPRTSTSLCTPGCKTGALMGCKNTATCNCSHVSK 57  
 CC 1 MSTRDPNLDVSVSKDGSAPRTSTSLCTPGCKTGALMGCKNTATCNCSHVSK 57  
 CC  
 CC RESULT 3  
 CC LANS\_BACSU STANDARD; PRT; 56 AA.  
 CC AC P10946;  
 CC DT 01-JUL-1989 (Rel. 11, Created)  
 CC DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Lantibiotic subtilin precursor.  
 CC GN SPAS OR SUB.  
 CC OS Bacillus subtilis.  
 CC OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC OX NCBI\_TaxID=1423;  
 CC RX MEDLINE=88243844; PubMed=2837490;  
 CC RA Banerjee S., Hansen J.N.;  
 CC RT "Structure and expression of a gene encoding the precursor of  
 CC subtilin, a small protein antilipidic."  
 CC RL J Biol. Chem. 263:9508-9514(1988).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN-ATCC 6633 / LH45;  
 CC MEDLINE=92138640; PubMed=1735728;  
 CC

RA Chung Y.J., Steen M.T., Hansen J.N.;  
 RT "The subtilin gene of Bacillus subtilis ATCC 6633 is encoded in an  
 RT operon that contains a homolog of the hemolysin B transport  
 RT protein."  
 RT J. Bacteriol. 174:1417-1422(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 6633 / LH45;  
 RX MEDLINE=92171481; PubMed=1539969;  
 RA Klein C., Kaletta C., Schnell N., Entian K.-D.;  
 RT "Analysis of genes involved in biosynthesis of the lantibiotic  
 RT subtilin."  
 RL Appl. Environ. Microbiol. 58:132-142(1992).  
 RN [4]  
 RP SEQUENCE OF 25-56.  
 RX MEDLINE=75040028; PubMed=4154277;  
 RA Gross E., Kiltz H.H., Nebelin E.;  
 RT "Subtilin, VI: the structure of subtilin."  
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:810-812(1973).  
 RN [5]  
 RP MODE OF ACTION.  
 RX MEDLINE=89276381; PubMed=2471644;  
 RA Schueller F., Benz R., Sahl H.-G.;  
 RT "The peptide antibiotic subtilin acts by formation of  
 RT voltage-dependent multi-state pores in bacterial and artificial  
 RT membranes."  
 RL Eur. J. Biochem. 182:181-186(1989).  
 RN [6]  
 RP STRUCTURE BY NMR.  
 RX STRAIN-ATCC 6633 / LH45;  
 RX MEDLINE=92192284; PubMed=1547888;  
 RA Chan W.C., Bycroft B.W., Leylands M.L., Llan L.-Y., Yang J.C.,  
 RA Roberts G.C.K.;  
 RT "Sequence-specific resonance assignment and conformational analysis  
 RT of subtilin by 2D NMR."  
 RL FEBS Lett. 300:56-62(1992).  
 RN [7]  
 RP MUTAGENESIS OF SER-29.  
 RX MEDLINE=93167832; PubMed=8434932;  
 RA Liu W., Hansen J.N.;  
 RT "The antimicrobial effect of a structural variant of subtilin against  
 RT outgrowing Bacillus cereus T spores and vegetative cells occurs by  
 RT different mechanisms."  
 RL Appl. Environ. Microbiol. 59:648-651(1993).  
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)  
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL  
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS  
 CC TRANSMEMBRANE PORES.  
 CC -1- P.T.M.: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF  
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE  
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND  
 CC CLEAVAGE OF THE MODIFIED PRECURSOR.  
 CC -1- MISCELLANEOUS: SUBTILIN ACTIVITY IS OBSERVED DURING STATIONARY  
 CC PHASE, BUT NOT DURING EXPONENTIAL GROWTH.  
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER LANTIBIOTICS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J03767; AAA22841.1; -  
 CC EMBL: M86869; AAA22840.1; -  
 CC EMBL: M83944; AAA22772.1; -  
 CC EMBL: M99263; AAA22778.1; -  
 CC EMBL: U09819; AAB91589.1; -  
 CC PIR: A28112; NTBSSA.  
 CC InterPro: IPR001049; Gallidermln.  
 CC InterPro: IPR000446; Nisin.

```

DR Pfam: PF02052; Gallidermin; 1.
DR PRINTS; PR00324; NISIN.
KW Antibiotic; Bacteriocin; Lantibiotic.
FT PROPEP 1 24
FT MOD_RES 25 56 LANTIBIOTIC SUBTILIN.
FT MOD_RES 27 27 D-ALANINE.
FT MOD_RES 29 29 DHA (2,3-DIDEHYDROALANINE).
FT MOD_RES 32 32 D-ABU (AMINOBTYRIC ACID).
FT MOD_RES 37 37 D-ABU (AMINOBTYRIC ACID).
FT MOD_RES 42 42 DHB (2,3-DIDEHYDROBTYRINE).
FT MOD_RES 47 47 D-ABU (AMINOBTYRIC ACID).
FT MOD_RES 49 49 D-ABU (AMINOBTYRIC ACID).
FT MOD_RES 55 55 DHA (2,3-DIDEHYDROALANINE).
FT THIOETH 27 31 ALA-S-CYS (LANTHIONINE).
FT THIOETH 32 35 ABU-S-CYS (BETA-METHYLLANTHIONINE).
FT THIOETH 37 43 ABU-S-CYS (BETA-METHYLLANTHIONINE).
FT THIOETH 47 50 ABU-S-CYS (BETA-METHYLLANTHIONINE).
FT THIOETH 49 52 ABU-S-CYS (BETA-METHYLLANTHIONINE).
FT MUTAGEN 29 29 S-2A: DEVOID OF ANTIMICROBIAL ACTIVITY;
KEEPS FULL LYSIS CAPACITY.
SEQUENCE 56 AA: 6218 MW; DA9707BF8A1EBBA CRC64;
Query Match 51.6%; Score 157; DB 1; Length 56;
Best Local Similarity 59.2%; Pred. No. 5.5e-13;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

OY 5 FNLDIVSVSKKDSGASPRITSTSLCTPGCKTGALMGCKMKTANCHRSI 53
DB 6 DFDDIVVKSVDKSKITPOMKSESLCTPGCVTGALQCFQTLQCNCKI 54

RESULT 4
LANE_STAGP STANDARD; PRT: 52 AA.
ID LANE_STAGP Q54093;
AC P08136; Q54093;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lantibiotic epidermin precursor.
GN EPTA.
OS Staphylococcus epidermidis.
OG Plasmid ptu 32.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TU 3298 / DSM 3095;
RC MEDLINE=88216821; PubMed=2835685;
RC Schnell N., Entian K.-D., Schneider U., Gotz F., Zahner H.,
RC Kellner R., Jung G.;
RC "Prepeptide sequence of epidermin, a ribosomally synthesized
RT antibiotic with four sulphide-rings.";
RL Nature 333:276-278(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TU 3298 / DSM 3095;
RC MEDLINE=92155237; PubMed=1740156;
RC Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RC Goetz F., Entian K.-D.;
RC "Analysis of genes involved in the biosynthesis of lantibiotic
RT epidermin.";
RL Eur. J. Biochem. 204:57-68(1992).
CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
CC TRANSMEMBRANE PORES.
CC -1- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF
CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND
CC CLEAVAGE OF THE MODIFIED PRECURSOR.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X07840; CA30689.1; -
DR EMBL: X07840; CA30690.1; -
DR EMBL: X62386; CA44252.1; -
DR EMBL: A12927; CAA01070.1; -
DR PIR: S00768; EPSED.
DR InterPro: IPR001049; Gallidermin.
DR Pfam: PF02052; Gallidermin; 1.
DR PRINTS; PR00323; GALLIDERMIN.
KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.
FT PROPEP 1 30
FT CHAIN 31 52 LANTIBIOTIC EPIDERMIN.
FT MOD_RES 33 33 D-ALANINE.
FT MOD_RES 38 38 D-ABU (AMINOBTYRIC ACID).
FT MOD_RES 44 44 DHB (2,3-DIDEHYDROBTYRINE).
FT MOD_RES 46 46 D-ALANINE.
FT MOD_RES 49 49 DHA (2,3-DIDEHYDROALANINE).
FT MOD_RES 52 52 S-(2-AMINOVINYL)-D-CYSTEINE.
FT THIOETH 33 37 ALA-S-CYS (LANTHIONINE).
FT THIOETH 38 41 ALA-S-CYS (BETA-METHYLLANTHIONINE).
FT THIOETH 46 51 ALA-S-CYS (LANTHIONINE).
FT THIOETH 49 52 DHA-S-CYS (AVI).
FT THIOETH 52 52 DHA-S-CYS (AVI).
SEQUENCE 52 AA: 5632 MW; 8B1AD2875BF1D6D CRC64;
Query Match 31.1%; Score 94.5; DB 1; Length 52;
Best Local Similarity 62.2%; Pred. No. 2.7e-05;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

OY 6 FNLDIVSVSKK---DSGASPRITSTSLCTPGC-KTGA 38
DB 11 FNLD-VKVNKESNDSGAPRIASKFTCTPGCAKTGS 46

RESULT 5
LANG_STAGA STANDARD; PRT: 52 AA.
ID LANG_STAGA P21838;
AC P21838;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lantibiotic gallidermin precursor.
GN GDMA.
OS Staphylococcus gallinarum.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1293;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=89305540; PubMed=2765032;
RC Schnell N., Entian K.-D., Goetz F., Hoerner T., Kellner R., Jung G.;
RC "Structural gene isolation and prepeptide sequence of gallidermin, a
RT new lantithionine containing antibiotic.";
RL FEMS Microbiol. Lett. 49:263-267(1989).
RN [2]
RP SEQUENCE OF 31-52.
RC STRAIN-TU 3928;
RC MEDLINE=89030695; PubMed=3181159;
RC Kellner R., Jung G., Hoerner T., Zaehner H., Schnell N., Entian K.-D.,
RC Goetz F.;
RC "Gallidermin: a new lantithionine-containing polypeptide antibiotic.";
RL Eur. J. Biochem. 177:53-59(1988).
RN [3]
RP STRUCTURE BY NMR.
RC MEDLINE=92032577; PubMed=1932575;
RC Freund S., Jung G., Gutbrod O., Folkers G., Gibbons W.A., Allgauer H.,
RC Werner R.;
RC "The solution structure of the lantibiotic gallidermin.";

```

BL001ymers 31:803-811(1991).

-1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC) ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS TRANSMEMBRANE PORES.

-1- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND CLEAVAGE OF THE MODIFIED PRECURSOR.

-1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.

PIR: A61072; EPCSD.

InterPro: IPR001049; GalldermIn.

DR Pfam: PF02052; GalldermIn.1.

DR PRINTS: PR00323; GALDERMIN.

KW Antibiotic; Bacteriocin; Lantibiotic.

FT PROPEP 1 30

FT CHAIN 1 30

FT MOD\_RES 31 52 LANTIBIOTIC GALLDERMIN.

FT MOD\_RES 33 33 D-ALANINE.

FT MOD\_RES 38 38 D-ABU (AMINOBUTYRIC ACID).

FT MOD\_RES 44 44 DHB (2,3-DIDEHYDROBUTYRINE).

FT MOD\_RES 46 46 D-ALANINE.

FT MOD\_RES 49 49 DHA (2,3-DIDEHYDROALANINE).

FT MOD\_RES 52 52 S-(2-AMINOVINYL)-D-CYSTEINE.

FT THIOETH 33 37 ALA-S-CYS (LANTHIONINE).

FT THIOETH 38 41 ABU-S-CYS (BETA-METHYLLANTHIONINE).

FT THIOETH 46 51 ALA-S-CYS (LANTHIONINE).

FT THIOETH 49 52 DHA-S-CYS (AVI).

FT THIOETH 52 52

SO SEQUENCE 52 AA; 5647 MW; 8584C0040AB4786D CRC64;

Query Match 30.1%; Score 91.5; DB 1; Length 52;

Best Local Similarity 62.2%; Pred. No. 6.4e-05;

Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

QY 6 FNLDLVSVSKR---DSGASPRITSTSLCTPGC-KTGA 38

DB 11 FDLID-VKVNKESDSEAFRIASKEFCTGCATKTS 46

RESULT 6

SRTA\_STRPY STANDARD; PRT; 46 AA.

AC 09FDVI;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update).

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lantibiotic str. precursor.

SRTA OR SPY1083.

Streptococcus pyogenes.

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

CC Streptococcus.

OX NCBI\_Taxid-1314;

CC [1]

SEQUENCE FROM N.A.

RP STRAIN-BL-T;

RC STRAIN-BL-T;

RA Karaya K., Takeo A.;

RT "Gene cluster of lantibiotics producing by Streptococcus pyogenes."

RL Submitted (MUG-1999) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.

RC STRAIN-SF370 / ATCC 700294 / Serotype M1;

RX MEDLINE-11192684; PubMed-11296296;

RA Perrelet J.J., McShan W.W., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Szatec S., Suvorov A.N., Kenon S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;

RT Complete genome sequence of an M1 strain of Streptococcus pyogenes.;

RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AB030831; BAB08162.1;

DR EMBL: AE006552; AAK33966.1;

KW Antibiotic; Bacteriocin; Lantibiotic; Complete proteome.

FT PROPEP 1 46

FT CHAIN 1 46

SO SEQUENCE 46 AA; 5219 MW; 3775C548BA2800F CRC64;

Query Match 23.7%; Score 72; DB 1; Length 46;

Best Local Similarity 56.7%; Pred. No. 0.015;

Matches 17; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

QY 4 KDFNLDLVSVSKKDSGASPRITSTSLCTPG 33

DB 6 KDFDLID-KTNKKDT-ATPYGSRYLCTPG 33

RESULT 7

Y109\_YEAST STANDARD; PRT; 995 AA.

AC P40442;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-FEB-2002 (Rel. 41, Last annotation update)

DE Hypothetical 99.7 kDa protein in Sdl1 5'region precursor.

GN Y1169C OR Y19402.07C.

OS Saccharomyces cerevisiae (Baker's Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_Taxid-4932;

CC [1]

SEQUENCE FROM N.A.

RP STRAIN-6288C / AB972;

RC Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D., Church C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A., Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M., Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D., Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V., Walsh S.V., Whitehead S.;

RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z46921; CAAB7023.1;

DR SGD: S0001431; Y1169C.

DR InterPro: IPR004089; Chmtaxs\_transd.

DR InterPro: IPR000427; T\_SNARE.

DR PROSITE: PS50192; T\_SNARE; UNKNOWN1.

KW Hypothetical protein; Signal.

FT CHAIN 1 23

FT DOMAIN 1 23

FT CAROHD 92 95

FT CAROHD 92 154

FT CAROHD 28 28

FT CAROHD 35 35

FT CAROHD 468 468

FT CAROHD 664 664

FT CAROHD 664 664

SO SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match 20.7%; Score 63; DB 1; Length 995;

Best Local Similarity 26.9%; Pred. No. 4.3;

Matches 21; Conservative 8; Mismatches 25; Indels 24; Gaps 2;

QY 3 TKDFNLDLVSVSKKDSGASPRITSTSLCTPGC-KTGA 40

Db 676 TTDSNGNVYTTTTTTPCSTTATITSCDETCHYSTIGAVTIEVSSKSTTATVTHCD 735

Qy 41 --GCNMKTATCHCSIHS 56

Db 736 DNGCNKKTVTSSECKETS 753

RESULT 8  
YF48\_HUMAN

ID	YEAR	HUMAN	STANDARD	PRT	132 AA
AC	09H9C44	Q9H975			
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Hypothetical protein KIAA1548.				
GN	KIAA1548				
GN	Homo sapiens (human).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiina; Homo.				
NCBI	TaxID:9606;				
OR					

[1]  
 SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshitaka Y.,  
 RA Maksumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahari K., Masuno Y., Sasaki T.,  
 RT "NDO human cDNA sequencing project",  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBU databases.  
 RP [2]  
 SEQUENCE OF 188-732 FROM N.A.

CC  
-1. SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.  
RT  
code for large proteins in vitro."  
RT  
xviii. The complete sequences of 100 new cDNA clones from brain which  
RA  
"Prediction of the coding sequences of unidentified human genes.  
RA  
Nagase T., Kituno R., Nakayama M., Hirotsawa M., Ohara O.,  
RX  
MEDLINE=20450683; PubMed=1097877;  
RC  
TISSUE=Brain;  
RC  
TISSUE=Brain;

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL: AK023019; BAB14360.1; -  
DR EMBL: AB046768; BAB13374.1; -  
DR InterPro: IPR000299; Band\_4.1.  
DR Pfam, PF00373; Band\_4; 1.  
DR PRINTS: PR00935; BAND4.1.  
DR SMART; SM00295; B4; 1.  
DR PROSITE; PS00660; BAND\_4.1; 1.  
DR PROSITE; PS00661; BAND\_4.2; 1.  
DR PROSITE; PS50057; BAND\_4.3; 1.  
KW Hypothetical protein; Cytoskeleton.  
FT DOMAIN 40 235 BAND\_4.1-LIKE.  
FT CONFLICT 669 732  
FT TSSGPIAEAVYKQKCLITTEL -> LMSHGRSCSEAV

FT	SEQUENCE	732	AA:	81758	MM:	76D5BDBCCE09DE761	CRC64:	VF7D4 (IN REF. 1):
SO	Query Match	20.1%:	Score 61:	DB 1:	Length 732:			
	Best Local Similarity	29.4%:	Pred. No. 5.5:					
Matches	15;	Conservative	12;	Mismatches	16;	Indels	8;	Gaps 2

Qy 2 SYKDNLDLVSVSK--DSGSPRISTSLCTP-----GCKTALMGCMN 44  
::: - : ::- : : ::::: || | - : |||  
Db 631 ATDELALLASTENLIDHTVAPOVSTSIITRWIVPGASNSGLAGCEM 683

## RESULT 9

RINI_RAT	STANDARD;	PRT;	456 AA.
ID	RINI_RAT	.	

AC P29315; [View](#)  
DT 01-DEC-1992 (rel. 24, Created)  
DT 01-DEC-1992 (rel. 24, Last sequence update)  
DT 16-OCT-2001 (rel. 40, Last annotation update)  
DE Ribonuclease inhibitor.  
DE RNH.  
GN Rattus norvegicus (Rat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus  
OX NCBI\_TaxID=10116;

RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP [1]  
 RC TISSUE=Lung;  
 RX MEDLINE=92162755; PubMed=1536887;  
 RA Kanamomo M., Motojima K., Sasai M., Hattori H., Goto S.;  
 RT "cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue  
 distribution of the mRNA.";  
 RL Biochim. Biophys. Acta 1159:335-338(1992).  
 CC -I- FUNCTION: INHIBITOR OF PANCREATIC RNASE AND ANGIOGENIN. MAY ALSO  
 CC FUNCTION IN THE MODULATION OF CELLULAR ACTIVITIES.  
 CC -I- SUBUNIT: FORMS A TIGHT ONE-TO-ONE COMPLEX WITH THE RNASE.

CC - SUBCELLULAR LOCATION: Cytoplasm;c.  
CC - TISSUE SPECIFICITY: BRAIN, HEART, LUNG, LIVER, SPLEEN, TESTES AND  
CC KIDNEY; HIGHEST IN THE LUNG AND LOWEST IN THE HEART.  
CC - SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).  
CC

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC -----
DR EMBL, X62558; CAA44388.1; -.
DR DR PIR, S20597; S20597.
DR HSP, P10775; 2BNH.
DR Interpro: IPR001611; LRR.
DR Interpro: IPR003590; LRR_RNinh.
DR Interpro: IPR003592; LRR_out.
DR Pfam, PF00560; LRR_3.
DR SMART, SM00370; LRR_12.
DR SMART, SM00368; LRR_R1; 1.
KW Repeat, Leucine-rich repeat.  rep 31
```

FT	REPEAT	15	43	LRR A1.
FT	REPEAT	44	71	LRR B1.
FT	REPEAT	72	100	LRR A2.
FT	REPEAT	101	128	LRR B2.
FT	REPEAT	129	157	LRR A3.
FT	REPEAT	158	185	LRR B3.
FT	REPEAT	186	214	LRR A4.
FT	REPEAT	215	242	LRR B4.
FT	REPEAT	243	271	LRR A5.
FT	REPEAT	272	299	LRR B5.
FT	REPEAT	300	328	LRR A6.
FT	REPEAT	329	356	LRR B6.
FT	REPEAT	357	385	LRR A7.
FT	REPEAT	386	413	LRR B7.
FT	REPEAT	414	442	LRR A8.

SQ		SEQUENCE	436 AA;	49503 RM;	CGAGGACATCCTGGCCGCCGCCTTTC
Query Match	19.4%;	Score 59;	DB I;	Length 456;	
Best Local Similarity	34.0%;	Pred. No. 6;			
Matches 17; Conservative	8;	Mismatches 19;	Indels 6;	Gaps 3	

QY 8 LDIVASVKKDSGASPRITSTLCIPGKGTAL--MGCNMTATC--HOSI 53  
 Db 284 LSIAGNELDEGA-QLLCESTLEPGQLESIMWKTCSLTIAASCPHPCSV 333

## RESULT 10

MTCTU\_HELP0 STANDARD; PRT; 64 AA.  
 ID MTCTU\_HELP0  
 AC P55947;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Copper-metallothionein (Cu-MT).  
 OS Helix pomatia (Roman snail) (Edible snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 OC Helicacea; Helicidae; Helix.  
 OX NCBI\_TaxID=6536;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=9373947; PubMed=9230430;  
 RA Dallingner R., Berger R., Hunziker P.E., Kaegi J.H.R.;  
 RT "Metallothionein in snail Cd and Cu metabolism."  
 RL Nature 388:237-238(1997).  
 CC -1- FUNCTION: THE METALLOTHIONEIN IS INVOLVED IN THE CELLULAR  
 SEQUESTRATION OF TOXIC METAL IONS AND REGULATION OF ESSENTIAL  
 TRACE ELEMENTS. THIS ISOFORM BINDS EXCLUSIVELY COPPER.  
 CC -1- DOMAIN: 14 CYSTEINE RESIDUES ARE ARRANGED IN C-X-C GROUPS. THESE  
 ARE THOUGHT TO BE THE METAL-BINDING SITES IN OTHER  
 METALLOTHIONEINS.  
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.  
 CC HSP; P05106; IUV2.  
 KW Metal-binding; Metal-cholate cluster; Copper; Acetylation.  
 DR MOD\_RES 1 7  
 FT METAL 11 11 COPPER.  
 FT METAL 16 16 COPPER.  
 FT METAL 18 18 COPPER.  
 FT METAL 22 22 COPPER.  
 FT METAL 24 24 COPPER.  
 FT METAL 28 28 COPPER.  
 FT METAL 30 30 COPPER.  
 FT METAL 33 33 COPPER.  
 FT METAL 36 36 COPPER.  
 FT METAL 38 38 COPPER.  
 FT METAL 43 43 COPPER.  
 FT METAL 45 45 COPPER.  
 FT METAL 49 49 COPPER.  
 FT METAL 55 55 COPPER.  
 FT METAL 57 57 COPPER.  
 FT METAL 61 61 COPPER.  
 FT METAL 63 63 COPPER.  
 SO SEQUENCE 64 AA; 6205 MW; 96CC1998B7E12297 CRC64;  
 Key Match 19.2%; Score 58.5; DB 1; Length 64;  
 Best Local Similarity 50.0%; Pred. No. 0.96;  
 Matches 12; Conservative 3; Mismatches 6; Indels 3; Gaps 2;  
 Oy 30 CTPGCKTALMGCM-KTATCHCS 52  
 Db 18 CGNDCKCGA--GCNDRCCSCHCS 39  
 RESULT 11  
 TLE3\_HUMAN  
 ID TLE3\_HUMAN STANDARD; PRT; 772 AA.  
 AC 004726; Q9HGM5; O8WVR2;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Transducin-like enhancer protein 3 (ESG3).  
 GN TLE3 OR KIA1547.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC Tissue=fetal brain;

RX MEDLINE=93265135; PubMed=1303260;  
 RA Stifani S., Blumheller C.M., Redhead N.J., Hill R.E.,  
 RA Artavanis-Tsakonas S.;  
 RT "Human homologs of a Drosophila Enhancer of split gene product define  
 RT a novel family of nuclear proteins.";  
 RL Nat. Genet. 2:119-127(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=20450683; PubMed=10997877;  
 RA Nagase T., Kikuno R., Nakayama M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes.  
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which  
 RT code for large proteins in vitro."  
 RL DNA Res. 7:273-281(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Pancreas;  
 RA Strausberg R.;  
 RL Submitted (JCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: NUCLEAR EFFECTOR MOLECULE.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Placenta and lung.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TTP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE GRCOCHO/TLE FAMILY OF WD-REPEAT  
 PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; M99438; AAA61194.1; -  
 CC DR EMBL; AB046767; BAB13373.1; ALT\_INIT.  
 CC DR EMBL; BC015729; AAH15729.1; -  
 CC DR Genew; HGNC:11839; TLE3.  
 CC DR MIM; 600190; -  
 CC DR InterPro; IPR001680; WD40.  
 CC DR Pfam; PF00400; WD40; 6.  
 CC DR PRINTS; PR00320; GPROTEINRPT.  
 CC DR Prodom; PD000018; WD40; 1.  
 CC DR SMART; SM00320; WD40; 7.  
 CC DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
 CC DR PROSITE; PS50082; WD\_REPEATS\_2; 2.  
 CC DR PROSITE; PS50294; WD\_REPEATS\_REGION; 2.  
 CC KW Nuclear protein; Phosphorylation; Repeat; WD repeat;  
 CC KW Alternative splicing.  
 CC FT DOMAIN 1 131 GLN-RICH.  
 CC FT DOMAIN 122 198 GLY/PRO-RICH.  
 CC FT DOMAIN 199 268 CCN DOMAIN.  
 CC FT DOMAIN 269 451 SER/PRO-RICH.  
 CC FT DOMAIN 225 228 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC FT DOMAIN 401 409 POLY-ALA.  
 CC FT REPEAT 484 522 WD 1.  
 CC FT REPEAT 530 569 WD 2.  
 CC FT REPEAT 574 613 WD 3.  
 CC FT REPEAT 616 655 WD 4.  
 CC FT REPEAT 657 696 WD 5.  
 CC FT REPEAT 698 737 WD 6.  
 CC FT REPEAT 739 771 WD 7.  
 CC FT MOD\_RES 240 240 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
 CC FT MOD\_RES 229 259 PHOSPHORYLATION (BY CDC2) (POTENTIAL).  
 CC FT MOD\_RES 263 263 PHOSPHORYLATION (BY CDC2) (POTENTIAL).  
 CC FT MOD\_RES 267 267 PHOSPHORYLATION (BY CDC2) (POTENTIAL).  
 CC FT VARSPPLIC 342 353 MISSING (IN ISOFORM 2).  
 CC FT VARSPPLIC 351 353 MISSING (IN ISOFORM 3).  
 CC FT VARSPPLIC 417 421 MISSING (IN ISOFORM 3).  
 CC FT CONFLICT 229 229 A -> V (IN REF. 1).

```

FT CONFLICT 487 487 E -> G (IN REF. 1).
FT CONFLICT 498 498 T -> S (IN REF. 1).
FT CONFLICT 535 535 I -> M (IN REF. 1).
FT CONFLICT 541 541 L -> H (IN REF. 1).
FT CONFLICT 553 553 A -> G (IN REF. 1).
FT CONFLICT 692 692 D -> H (IN REF. 1).
FT CONFLICT 736 736 F -> S (IN REF. 1).
SQ SEQUENCE 772 AA; 83416 MW; A2A469D73BF04A43 CRC64;

Query Match 19.2%; Score 58.5; DB 1; Length 772;
Best Local Similarity 38.3%; Pred. No. 12;
Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

QY 2 SKRDFNLVSVSKSDSGASPRITSLCTPCKTGALMGCMKRTAT 48
Db 267 SPENGLDKARSLKRDAPTSVAVSSSTPSSKTKDL-GHNDKST 312

RESULT 12
KPGA_RAT STANDARD; PRT; 727 AA.
P51556;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Diacylglycerol kinase, alpha (EC 2.7.1.107) (Dilyceride kinase) (DGK-
alpha) (DAG kinase alpha) (80 kDa diacylglycerol kinase).
DGA OR DAGK1 OR DAGK.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
Tissue-Brain;
MEDLINE=93095720; PubMed=1339302;
Goto K., Watanabe M., Kondo H., Yusa H., Sakane F., Kanoh H.;
"Gene cloning, sequence, expression and in situ localization of 80
kDa diacylglycerol kinase specific to oligodendrocyte of rat brain.";
Brain Res. Mol. Brain Res. 16:75-87(1992).
-1- FUNCTION: UPON CELL STIMULATION CONVERTS THE SECOND MESSENGER
DIACYLGLYCEROL INTO PHOSPHATIDATE, INITIATING THE RESYNTHESIS
OF PHOSPHATIDYLINOSITOLS AND ATTENUATING PROTEIN KINASE C
ACTIVITY.
-1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-
diacylglycerol 3-phosphate.
-1- ENZYME REGULATION: STIMULATED BY CALCIUM AND PHOSPHATIDYLSELINE.
-1- PHOSPHORYLATED BY PROTEIN KINASE C (BY SIMILARITY).
-1- SUBUNIT: MONOMER (BY SIMILARITY).
-1- TISSUE SPECIFICITY: LYMPHOCTES AND OLIGODENDROGLIAL CELLS.
-1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
FAMILY.
-1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAINS.
-1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; S49760; AAB24434.1;
InterPro: IPR000756; DAGK.
InterPro: IPR001206; DAGK.
InterPro: IPR002219; DAG_PE-bind.
InterPro: IPR003622; DAG_Kin-cat.
InterPro: IPR002048; EF-hand.
Pfam: PF00036; ehand; 2.
Pfam: PF00130; DAG_PE-bind; 2.
Pfam: PF00609; DAGK; 1.

```

```

DR Pfam: PF00781; DAGK; 1.
DR PRINTS: PR00008; DAGEDOMAIN.
DR ProDom: PD000012; EF-hand; 1.
DR ProDom: PD002939; DAGK; 1.
DR ProDom: PD005043; DAG_Kin-cat; 1.
DR SMART: SM00109; C1; 2.
DR SMART: SM00045; DAGK; 1.
DR SMART: SM00046; DAGK; 1.
DR SMART: SM00054; Efh; 2.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.
DR PROSITE: PS00081; DAG_PE_BIND_DOM_2; 2.
DR PROSITE: PS00018; EF_HAND; 2.
KW Transferrase; kinase; Calcium-binding; Phorbol-ester binding;
KW Repeat; Multigene family.
FT CA_BIND 121 132 EF-HAND 1 (PROBABLE).
FT CA_BIND 166 177 EF-HAND 2 (PROBABLE).
FT DOMAIN 204 251 PHORBOL-ESTER AND DAG BINDING 1.
FT DOMAIN 268 317 PHORBOL-ESTER AND DAG BINDING 2.
FT DOMAIN 367 492 CATALYTIC-A (POTENTIAL).
FT DOMAIN 512 693 CATALYTIC-B (POTENTIAL).
SQ SEQUENCE 727 AA; 82198 MW; B5A248AD2F61C1D CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 727;
Best Local Similarity 24.6%; Pred. No. 15;
Matches 16; Conservative 6; Mismatches 28; Indels 15; Gaps 2;

QY 1 MSTRDFNLVSVSKSDSGASPRITSLCTPCKTGALMGCMKRTAT-----C 49
Db 246 MKRQPEVSTYAKSRNDIVQPHVWVRG---GHSGRCDRCOKIRFVHSILGLHCWC 301
QY 50 HCSIH 54
Db 302 HLEIH 306

RESULT 13
TLE3_MOUSE STANDARD; PRT; 771 AA.
ID TLE3_MOUSE
AC 008122;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transducin-like enhancer protein 3 (ESG).
GN TLE3 OR ESG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RC STRAIN=CD-1; PubMed=8365415;
RA MEDLINE=93373944; PubMed=8365415;
RA Miyasaka H., Choudhury B.K., Hou E.W., Li S.S.-L.;
"Gene cloning and expression of mouse and human cDNA encoding
RT AEG and ESG proteins with strong similarity to Drosophila enhancer of
RT split groucho protein.";
RT Eur. J. Biochem. 216:343-352(1993).
-1- FUNCTION: NUCLEAR EFFECTOR MOLECULE (BY SIMILARITY). MAY PLAY AN
IMPORTANT ROLE DURING SPERMATOGENESIS.
-1- SUBCELLULAR LOCATION: Nuclear (Probable).
-1- TISSUE SPECIFICITY: Expressed only in testis.
-1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
-1- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

DR EMBL: X73360; CA51770.1; -  
 DR PIR: S34162; S34162.  
 DR PIR: S35681; S35681.  
 DR MGI: 104634; Tle3.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 2.  
 DR NCBI: protein; Phosphorylation; Repeat; WD repeat; Spermatogenesis.  
 FT DOMAIN 1 130 GLN-RICH.  
 FT DOMAIN 131 197 GLY/PRO-RICH.  
 FT DOMAIN 198 267 CCN DOMAIN.  
 FT DOMAIN 268 450 SER/PRO-RICH.  
 FT DOMAIN 224 227 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT REPEAT 483 521 WD 1.  
 FT REPEAT 529 568 WD 2.  
 FT REPEAT 573 612 WD 3.  
 FT REPEAT 615 654 WD 4.  
 FT REPEAT 656 695 WD 5.  
 FT REPEAT 697 736 WD 6.  
 FT REPEAT 738 770 WD 7.  
 FT DOMAIN 400 408 POLY-ALA.  
 FT MOD\_RES 239 239 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
 FT MOD\_RES 258 258 PHOSPHORYLATION (BY CDC2) (POTENTIAL).  
 FT MOD\_RES 262 262 PHOSPHORYLATION (BY CDC2) (POTENTIAL).  
 FT MOD\_RES 266 266 PHOSPHORYLATION (BY CDC2) (POTENTIAL).  
 SO SEQUENCE 771 AA; 83113 MW; 428C9127EAF66DD CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 771;  
 Best Local Similarity 38.3%; Pred. No. 16;  
 Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

OY 2 STRFDLVLVSVSKDGSAPRTTSLCTPCKTALMGCMKKTAT 48  
 266 SEPPENGLDARGLKDKAPTSPASVSSSTPSKTKDL-GHNDKST 311

RESULT 14  
 ENTK\_MOUSE STANDARD; PRT; 1069. AA.  
 ID ENTK\_MOUSE P97435;  
 AC 01-NOV-1997 (rel. 35, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 15-JUN-2002 (rel. 41, Last annotation update)  
 Enterokinase (EC 3.4.21.9) (Enterokinase).  
 PRS7 OR ENTK.  
 Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX MEDLINE=98147142; PUBMED=9486188;  
 RA Yuan X., Zheng X., Lu D., Rubin D.C., Pung C.Y.M., Sadler J.E.;  
 RT "Structure of murine enterokinase (enteropeptidase) and expression in  
 small intestine during development."  
 RL Am. J. Physiol. 274:G342-G349(1998).  
 CC -1- FUNCTION: RESPONSIBLE FOR INITIATING ACTIVATION OF PANCREATIC  
 PROTEOLYTIC PHENOMENAS (TRYPsin, CHYMOTrypsin AND CARBOXYPEPTIDASE  
 A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN  
 TURN ACTIVATES OTHER PHENOMENAS INCLUDING CHYMOTRYPSINOGEN,  
 PROCARBOXYPEPTIDASES, AND PROELASTASES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-1-Ile-7 bond in  
 trypsinogen.  
 CC -1- SUBUNIT: HETERO-DIMER OF A CATALYTIC (LIGHT) CHAIN AND A  
 MULTIDOMAIN (HEAVY) CHAIN LINKED BY A DISULFIDE BOND (BY  
 SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).

CC -1- PTM: THE CHAINS ARE DERIVED FROM A SINGLE PRECURSOR THAT IS  
 CC CLEAVED BY A TRYPSIN-LIKE PROTEASE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SEA DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 CC EMBL: U73378; AAB37317.1; -  
 CC HSSP: Q07954; ICR8.  
 CC MEROPS: S01.156; -  
 CC MGI: 1197523; Prss7.  
 CC InterPro: IPR000859; CUB\_domain.  
 CC InterPro: IPR001314; Chymotrypsin.  
 CC InterPro: IPR002172; LDL\_recept\_A.  
 CC InterPro: IPR000998; MAM\_domain.  
 CC InterPro: IPR000082; Ser.protease\_Try.  
 CC InterPro: IPR001254; Srcr\_receptot.  
 CC InterPro: IPR001190; Srcr\_receptot.  
 CC Pfam: PF00057; Idl\_recept\_a; 2.  
 CC Pfam: PF00089; trypsin; 1.  
 CC Pfam: PF00431; CUB; 2.  
 CC Pfam: PF00530; SRCR; 1.  
 CC Pfam: PF00629; MAM; 1.  
 CC Pfam: PF01390; SEA; 1.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC SMART: SM00042; CUB; 2.  
 CC SMART: SM00192; LDLa; 2.  
 CC SMART: SM00137; MAM; 1.  
 CC SMART: SM00200; SEA; 1.  
 CC SMART: SM00202; SR; 1.  
 CC PROSITE: PS01180; CUB; 2.  
 CC PROSITE: PS01209; LDLA\_1; 2.  
 CC PROSITE: PS50068; LDLA\_2; 2.  
 CC PROSITE: PS00740; MAM\_1; 1.  
 CC PROSITE: PS50060; MAM\_2; 1.  
 CC PROSITE: PS50024; SEA; 1.  
 CC PROSITE: PS00420; SRCR\_1; FALSE\_NEG.  
 CC PROSITE: PS50287; SRCR\_2; 1.  
 CC PROSITE: PS50240; TRYPSIN\_DOM; 1.  
 CC PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
 CC SignalAnchor: Glycoprotein; Hydrolyase; Serine protease; Zymogen;  
 CC Transmembrane; Repeat.  
 CC CHAIN 1 829  
 CC CHAIN 830 1069  
 CC TRANSMEM 1 18  
 CC TRANSMEM 19 47  
 CC DOMAIN 48 1069  
 CC DOMAIN 52 169  
 CC DOMAIN 227 268  
 CC DOMAIN 270 379  
 CC DOMAIN 387 549  
 CC DOMAIN 569 679  
 CC DOMAIN 686 724  
 CC DOMAIN 723 816  
 CC DOMAIN 830 1069  
 CC ACT\_SITE 874 874  
 CC ACT\_SITE 925 925  
 CC ACT\_SITE 1021 1021  
 CC DISULFID 229 242  
 CC DISULFID 236 255  
 CC -----  
 CC NON-CATALYTIC CHAIN (HEAVY CHAIN).  
 CC CATALYTIC CHAIN (LIGHT CHAIN).  
 CC CYTOPLASMIC (POTENTIAL).  
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 CC (POTENTIAL).  
 CC EXTRACELLULAR (POTENTIAL).  
 CC SEA.  
 CC LDL-RECEPTOR CLASS A 1.  
 CC CUB 1.  
 CC MAM.  
 CC CUB 2.  
 CC LDL-RECEPTOR CLASS A 2.  
 CC SRCR.  
 CC SERINE PROTEASE.  
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC BY SIMILARITY.  
 CC BY SIMILARITY.



FT DISULFID 249 266 BY SIMILARITY.  
 FT DISULFID 688 700 BY SIMILARITY.  
 FT DISULFID 695 713 BY SIMILARITY.  
 FT DISULFID 707 722 BY SIMILARITY.  
 FT DISULFID 817 945 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 859 875 BY SIMILARITY.  
 FT DISULFID 959 1027 BY SIMILARITY.  
 FT DISULFID 991 1006 BY SIMILARITY.  
 FT DISULFID 1017 1045 BY SIMILARITY.  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 675 675 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 770 770 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 791 791 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 897 897 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 936 936 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 999 999 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1069 AA; 118735 MW; E62549E463743C3D CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 1069;

Best Local Similarity 39.3%; Pred. No. 22; Mismatches 24; Indels 7; Gaps 4;

Matches 24; Conservative 6; Mismatches 24; Indels 7; Gaps 4;

OY 2 STDFENIDL-VSVSKKDGASPRITSTSL-CTPCKTGALMGCMKMTAT-CHCSITHV 55  
 DB 167 SLSDFTTAVPTSDKLTITSSPMTISLGNLSITVAATTSAPL-CNLSTATFATISGHV 225

OY 56 S 56  
 DB 226 S 226

RESULT 15

TRPF\_LIPST STANDARD; PRT: 232 AA.

ID TRPF\_LIPST

AC 001128;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE N-(5'-phosphoribosyl)anthranilate isomerase (EC 5.3.1.24) (PRAI).

TRP1.

Lipomyces starkeyi.

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Lipomycetaceae; Lipomyces.

NCBI\_TaxID=29829;

SEQUENCE FROM N.A.

STRAIN=NCYC 1436;

MEDLINE=96269934; PubMed=8662214;

Bignell G.R., Bruce I.J., Evans I.H.;

"Electrophoretic karyotype of the amylolytic yeast Lipomyces starkeyi

and cloning, sequencing and chromosomal localization of its TRP1

gene.";

Curr. Genet. 30:83-88(1996).

-1- CATALYTIC ACTIVITY: N-(5'-phospho-beta-D-riboseyl)-anthranilate - 1-

-1- PATHWAY: Tryptophan biosynthesis; third step.

-1- SIMILARITY: BELONGS TO THE TRPF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL: Z68292; CAA92584.1;  
 DR HSSP: Q56320; IDI3.  
 DR InterPro: IPR001240; PRAI.  
 DR Pfam: PF00697; PRAI; 1.  
 KW isomerase; Tryptophan biosynthesis.  
 SQ SEQUENCE 232 AA; 24625 MW; E6FF1B4EA7D7A9E0 CRC64;

Query Match 18.6%; Score 56.5; DB 1; Length 232;  
 Best Local Similarity 36.7%; Pred. No. 6.2; Mismatches 11; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

OY 24 ITSTSLCTPCKTGALMGCMKMTAT-CHCSI 53  
 DB 3 VSTSLCTPIVKI-----CGLTVEAHCAI 27

Search completed: June 7, 2003, 15:16:49  
 Job time : 13 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:13:09 ; Search time 81 seconds  
(without alignments)  
144.996 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304  
Sequence: 1 MSTKDFNLDSVSKKSGA.....ALMGCMKATNCCHSHSK 57

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 671580 seqs, 206047115 residues

1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

SPREMBL\_21.\*  
1: sp\_Archea:\*  
2: sp\_Bacteria:\*  
3: sp\_Fungi:\*  
4: sp\_Human:\*  
5: sp\_Invertebrate:\*  
6: sp\_Mammal:\*  
7: sp\_Mhc:\*  
8: sp\_Organelle:\*  
9: sp\_Phage:\*  
10: sp\_Plant:\*  
11: sp\_Rodent:\*  
12: sp\_Virus:\*  
13: sp\_Vertebrate:\*  
14: sp\_Unclassified:\*  
15: sp\_Virus:\*  
16: sp\_Bacteriap:\*  
17: sp\_Archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	155	51.0	56	2	093GH3
2	125	41.1	56	2	093GH5
3	67	22.0	456	11	091V17
4	63	20.7	567	4	08WUL3
5	63	20.7	1140	4	096KG7
6	62	20.4	752	10	09FKZ9
7	61.5	20.2	536	4	08TBS6
8	61	20.1	1352	5	09V5J7
9	60	19.7	456	11	0924P4
10	60	19.7	456	11	08T919
11	59.5	19.6	78	16	09PB60
12	59.5	19.6	330	5	018118
13	59.5	19.6	394	5	08SMW4
14	59.5	19.6	804	5	09YU48
15	59.5	19.6	1574	11	088281
16	59	19.4	115	4	09UG71

17	58.5	19.2	63	2	08VT57	08VT57 streptococ
18	58.5	19.2	65	5	095P49	095P49 helix pomat
19	58.5	19.2	178	5	09UAV9	09UAV9 caenorhabd
20	58.5	19.2	365	17	097B89	097B89 thermoplasm
21	58.5	19.2	843	15	090096	090096 human immun
22	58.5	19.2	846	15	091F99	091F99 human immun
23	58.5	19.2	1546	4	09NS27	09NS27 homo sapien
24	58.5	19.2	1546	4	075445	075445 homo sapien
25	58	19.1	565	10	09SIO6	09SIO6 arabidopsis
26	58	19.1	613	15	08USS4	08USS4 human immun
27	58	19.1	648	5	09NKD7	09NKD7 drosophila
28	58	19.1	689	13	08UVG4	08UVG4 raja erinac
29	58	19.1	701	5	09VJU4	09VJU4 drosophila
30	57.5	18.9	764	11	09JIT3	09JIT3 rattus norv
31	57	18.8	57	5	09N9H2	09N9H2 venterupis
32	57	18.8	59	5	09N9H1	09N9H1 ruditapes d
33	57	18.8	75	5	0901N5	0901N5 crassostrea
34	57	18.8	107	5	09NG19	09NG19 crassostrea
35	57	18.8	361	15	071036	071036 human immun
36	57	18.8	403	15	041582	041582 human immun
37	57	18.8	412	15	0802G0	0802G0 human immun
38	57	18.8	656	17	08T785	08T785 methanosarc
39	56.5	18.6	262	13	09PT79	09PT79 oryzias lat
40	56.5	18.6	852	15	070010	070010 human immun
41	56	18.4	66	5	09N628	09N628 conus catus
42	56	18.4	66	5	09N625	09N625 conus catus
43	56	18.4	66	5	09NCW2	09NCW2 conus catus
44	56	18.4	505	16	08YIO5	08YIO5 bruceella me
45	56	18.4	529	5	024331	024331 drosophila

## ALIGNMENTS

RESULT 1	093GH3	PRELIMINARY;	PRT;	56 AA.
ID	093GH3	PRELIMINARY;	PRT;	56 AA.
AC	093GH3;			
DT	01-DEC-2001 (TREMBLrel, 19, Created)			
DT	01-DEC-2001 (TREMBLrel, 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel, 21, Last annotation update)			
DE	Hypothetical 6.2 kda protein.			
GN	ERISB.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;			
OC	Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A13;			
RA	Stein T., Borchert S., Conrad B., Feesche J., Entian K.-D.,			
RA	Hofemeister J.,			
RT	"A subtilin-like gene cluster of Bacillus subtilis A13 encodes two			
RT	(putative) lantibiotics, ericin A and ericin S."			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF233755; AAL15569.1; -			
KW	Hypothetical protein.			
SO	SEQUENCE 56 AA; 6241 MW; DEDDAB0892A1EBBA CRC64;			
Query Match	51.0%; Score 155; DB 2; Length 56;			
Best Local Similarity	54.7%; Pred. No. 4.5e-13;			
Matches	29; Conservative 10; Mismatches 12; Indels 2; Gaps 1;			
OY	5 DENLDSVSKKSGASPRITSTSLCTGCKTGALMGCMKATATGCHSHSK 57			
DB	6 DFDLDVNVYSKDSKITPQWKSSEVCTPCGCVGLQTCFLQITTCNC--HISK 56			
RESULT 2	093GH5	PRELIMINARY;	PRT;	56 AA.
ID	093GH5	PRELIMINARY;	PRT;	56 AA.
AC	093GH5;			
DT	01-DEC-2001 (TREMBLrel, 19, Created)			



Query Match	20.1%	Score 61	DB 5	Length 1352
Best Local Similarity	38.7%	Pred. NO.	27	
Matches 12	Conservative 3	Mismatches 12	Indels 4	Gaps 1
Qy	30	CTPCKTGALMGCMKMTATCH-----CSIRHS 56		
Db	940	CLDGGDGLTGTCTSKPLSCHVLNNGIHAT 970		
RESULT 9				
0924P4				
ID	0924P4	PRELIMINARY:	PRT:	456 AA.
AC	01-DEC-2001 (TREMBLREL 19, Created)			
DT	01-DEC-2001 (TREMBLREL 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLREL 21, Last annotation update)			



```

Query Match      19.6%; Score 59.5; DB 5; Length 330;
Best Local Similarity 27.9%; Pred. No. 10;
Matches 12; Conservative 5; Mismatches 17; Indels 9; Gaps 1;

QY 23 RITSISLCTPGCKTGALMGCN-----WKATCHCSHYVS 56
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      275 RTTASQOQAPACSTSCNOSCNOPAMOCQPMQNSQCGOQYVS 317

RESULT 13
OBSWY4 PRELIMINARY; PRT; 394 AA.
AC OBSWY4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE RH68811.p.
GN CG17364.
OS Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
OX 11
RN RP
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Flise E.,
RA George R., Gonzalez M., Guerin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Parasas V., Park S.,
RA Patel S., Phouanenvong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Gelniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY094951; AM11304.1;
SQ SEQUENCE 394 AA; 40602 MW; 8D6ED85C06F3A3E1 CRC64;

Query Match      19.6%; Score 59.5; DB 5; Length 394;
Best Local Similarity 30.5%; Pred. No. 12;
Matches 18; Conservative 10; Mismatches 20; Indels 11; Gaps 1;

QY 10 LVSVKSKDSCGASPRITSTSLCTP-----GCKTGALMGCNKKATPCHCSHYVS 57
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|
      116 IAAVSKNSRASSRHTSGAGCSOHSRDDLDGCGGCGTSGSCRCSESTLSMNRHLOK 174

RESULT 14
OBSWY4 PRELIMINARY; PRT; 804 AA.
AC OBSWY4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG17364 protein.
GN CG17364.
OS Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
OX 11
RN RP
RC STRAIN-BERKELEY;
RA MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champagne M., Pfeiffer B.D.,
RA Wan H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

```

```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA De Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spter E., Spradling A.C., Stapleton M., Strong R.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003536; AAF49781.1;
DR FlyBase: FBgn0036391; CG17364.
DR InterPro: IPR000217; Tubulin.
DR PROSITE: PS00227; TUBULIN.
SQ SEQUENCE 804 AA; 85229 MW; BD57F6D931371345 CRC64;

Query Match      19.6%; Score 59.5; DB 5; Length 804;
Best Local Similarity 30.5%; Pred. No. 25;
Matches 18; Conservative 10; Mismatches 20; Indels 11; Gaps 1;

QY 10 LVSVKSKDSCGASPRITSTSLCTP-----GCKTGALMGCNKKATPCHCSHYVS 57
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|
      375 IAAVSKNSRASSRHTSGAGCSOHSRDDLDGCGGCGTSGSCRCSESTLSMNRHLOK 433

RESULT 15
OBSWY4 PRELIMINARY; PRT; 1574 AA.
AC OBSWY4;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MEGF6.
GN MEGF6.
OS Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_Taxid=10116;
OX 11
RN RP
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
RA MEDLINE-98360089; PubMed-9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RT like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL: AB011532; BAA32462.1;
DR HSSP: P00736; IAP0.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF_24.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00017; EGF_like; 19.
DR PROSITE: PS00010; ASX_HYDROXYL; 5.

```

Sat Jan 7 15:19:23 2003

us-10-082-618-5.rspt

Page 6

```

DR PROSITE: PS00022; EGF_1; UNKNOWN_23.
DR PROSITE: PS01865; EGF_2; 23.
DR PROSITE: PS01871; EGF_CA; 5.
RW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 1574 AA; 165445 MW; 2B48533D8F7F6E7 CRC64;

Query Match 19.6%; Score 59.5; DB 11; Length 1574;
Best Local Similarity 27.8%; Pred. No. 49;
Matches 15; Conservative 3; Mismatches 9; Indels 27; Gaps 2.

OY 23 RITSNSLCTPG-----CKTGALMGCMKMTATQHC 51
   1:1 1111 1:1 1111111
DB 757 RVTGSGCLCPKGTGDCGADCPGRGWGLGCGEIPCACEHGA--SCNPETGTCIC 808

```

Search completed: June 7, 2003, 15:18:31  
Job time : 97 secs .

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:16:28 ; Search time 14 Seconds  
(without alignments)  
119.793 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTRDFNLIVSVSKDSGA.....ALMGCNMTATCCHSIHVK 57

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCtus.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298	98.0	57	4	US-08-836-687B-20
2	295	97.0	57	4	US-08-836-687B-24
3	295	97.0	57	4	US-08-836-687B-26
4	292	96.1	57	4	US-08-836-687B-28
5	291	95.7	57	1	US-08-129-151A-2
6	291	95.7	57	2	US-08-715-579-2
7	291	95.7	57	4	US-08-836-687B-40
8	291	95.7	57	4	US-08-836-687B-42
9	289	95.1	57	4	US-08-836-687B-43
10	289	95.1	57	4	US-08-836-687B-46
11	288	94.7	57	4	US-08-836-687B-44
12	285	93.8	57	4	US-08-836-687B-45
13	285	93.8	83	3	US-08-773-711A-2
14	282.5	92.9	56	4	US-08-836-687B-41
15	280	92.1	57	4	US-08-836-687B-47
16	185	60.9	34	6	5231013-4
17	181	59.5	41	2	US-08-535-494-9
18	181	59.5	41	4	US-09-097-635-9
19	178	58.6	34	1	US-07-880-003-1
20	178	58.6	34	2	US-08-030-911-2
21	176	57.9	34	2	US-08-030-911-1
22	157	51.6	56	1	US-07-981-525-2
23	157	51.6	56	1	US-07-981-525-7
24	157	51.6	56	1	US-08-220-033-2
25	157	51.6	56	1	US-08-220-033-7
26	157	51.6	56	2	US-08-465-491-2
27	157	51.6	56	2	US-08-465-491-7

28	157	51.6	56	2	US-08-986-617-2	Sequence 2, Appl
29	157	51.6	56	2	US-08-986-617-7	Sequence 7, Appl
30	154	50.7	56	1	US-07-981-525-9	Sequence 9, Appl
31	154	50.7	56	1	US-08-220-033-9	Sequence 9, Appl
32	154	50.7	56	2	US-08-465-491-9	Sequence 9, Appl
33	154	50.7	56	2	US-08-986-617-9	Sequence 9, Appl
34	128	42.1	39	2	US-08-535-494-5	Sequence 5, Appl
35	128	42.1	39	4	US-09-097-635-5	Sequence 5, Appl
36	113	37.2	23	1	US-08-524-677-10	Sequence 10, Appl
37	113	37.2	23	2	US-08-465-491-25	Sequence 25, Appl
38	113	37.2	23	2	US-08-986-617-25	Sequence 25, Appl
39	106	34.9	34	4	US-08-836-687B-49	Sequence 49, Appl
40	105	34.5	34	4	US-08-836-687B-50	Sequence 50, Appl
41	105	34.5	34	4	US-08-836-687B-51	Sequence 51, Appl
42	104	34.2	34	2	US-08-392-625-15	Sequence 15, Appl
43	104	34.2	34	2	US-08-465-961A-15	Sequence 15, Appl
44	104	34.2	34	4	US-08-836-687B-48	Sequence 48, Appl
45	103	33.9	32	1	US-08-220-033-23	Sequence 23, Appl

## ALIGNMENTS

RESULT 1  
US-08-836-687B-20  
; Sequence 20, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/770  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-20

Query Match 98.0%; Score 298; DB 4; Length 57;  
Best Local Similarity 98.2%; Pred. No. 1.4e-29;  
Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTRDFNLIVSVSKDSGASPRITSLCTPCKGTGALMGCNMTATCCHSIHVK 57  
DB 1 MSTRDFNLIVSVSKDSGASPRITSLCTPCKGTGALMGCNMTATCCHSIHVK 57

RESULT 2  
US-08-836-687B-24  
; Sequence 24, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/770  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-24

Query Match 97.0%; Score 295; DB 4; Length 57;  
Best Local Similarity 96.5%; Pred. No. 3.3e-29;  
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57  
Db 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57

## RESULT 3

US-08-836-687B-26  
; Sequence 26, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-26

Query Match 97.0%; Score 295; DB 4; Length 57;  
Best Local Similarity 96.5%; Pred. No. 3.3e-29;  
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57  
Db 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57

## RESULT 4

US-08-836-687B-28  
; Sequence 28, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 28  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-28

Query Match 96.1%; Score 292; DB 4; Length 57;  
Best Local Similarity 94.7%; Pred. No. 7.6e-29;  
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57  
Db 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57

## RESULT 5

US-08-129-151A-2  
; Sequence 2, Application US/08129151A  
; Patent No. 5594103  
; GENERAL INFORMATION:  
; APPLICANT: DE VOS, Willem M.  
; APPLICANT: SIEZEN, Roelant J.  
; APPLICANT: KUIPERS, Oscar P.  
; TITLE OF INVENTION: LANTIBIOTICS SIMILAR TO NISIN A, LACTIC  
; TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR  
; TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR

; TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: C/O YOUNG & THOMPSON  
; STREET: 745 South 23rd Street, Second Floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/129,151A  
FILING DATE: 07-OCT-1993  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: NL 910634  
FILING DATE: 11-APR-1991

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/NL92/00068  
FILING DATE: 09-APR-1992

## ATTORNEY/AGENT INFORMATION:

NAME: PATCH, Andrew J.  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: BO 37078

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 703/521-2297  
TELEFAX: 703/685-0573

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 57 amino acids  
TYPE: amino acid

## TOPOLOGY: linear

## MOLECULE TYPE: protein

US-08-129-151A-2  
Query Match 95.7%; Score 291; DB 1; Length 57;  
Best Local Similarity 96.5%; Pred. No. 1e-28;  
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57  
Db 1 MSTKDNLDVSVSKDSGASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVS 57

## RESULT 6

US-08-715-579-2  
; Sequence 2, Application US/08715579  
; Patent No. 5928946  
; GENERAL INFORMATION:  
; APPLICANT: DE VOS, Willem M.  
; APPLICANT: SIEZEN, Roelant J.  
; APPLICANT: KUIPERS, Oscar P.  
; TITLE OF INVENTION: LANTIBIOTICS SIMILAR TO NISIN A, LACTIC  
; TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR  
; TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR  
; TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: C/O YOUNG & THOMPSON  
; STREET: 745 South 23rd Street, Second Floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



```

; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,579
; FILING DATE: 19-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/129,151
; FILING DATE: 07-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 9100634
; FILING DATE: 11-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00068
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: BO 37078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/521-2297
; TELEFAX: 703/685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-715-579-2

```

```

Query Match          95.7%: Score 291; DB 2; Length 57;
Best Local Similarity 96.5%: Pred. No. 1e-28;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

```

```

RESULT 7
US-08-836-687B-40
; Sequence 40, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 40
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
; US-08-836-687B-40

```

```

Query Match          95.7%: Score 291; DB 4; Length 57;
Best Local Similarity 96.5%: Pred. No. 1e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

```

```

RESULT 8
US-08-836-687B-43
; Sequence 43, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:

```

```

; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 43
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
; US-08-836-687B-43

```

```

Query Match          95.7%: Score 291; DB 4; Length 57;
Best Local Similarity 96.5%: Pred. No. 1e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

```

```

RESULT 9
US-08-836-687B-42
; Sequence 42, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 42
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
; US-08-836-687B-42

```

```

Query Match          95.1%: Score 289; DB 4; Length 57;
Best Local Similarity 96.5%: Pred. No. 1.8e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

```

```

RESULT 10
US-08-836-687B-46
; Sequence 46, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 46
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
; US-08-836-687B-46

```

```

Query Match          95.1%: Score 289; DB 4; Length 57;
Best Local Similarity 96.5%: Pred. No. 1.8e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57  
Db 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 11  
US-08-836-687B-44  
; Sequence 44, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 44  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-44

Query Match 94.7%; Score 288; DB 4; Length 57;  
Best Local Similarity 96.5%; Pred. No. 2.3e-28;  
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57  
Db 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 12  
US-08-836-687B-45  
; Sequence 45, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-45

Query Match 93.8%; Score 285; DB 4; Length 57;  
Best Local Similarity 94.7%; Pred. No. 5.4e-28;  
Matches 54; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57  
Db 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 13  
US-08-773-731A-2  
; Sequence 2, Application US/08773731A  
; Patent No. 6100056  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael J.  
; APPLICANT: Dodd, Helen M.  
; TITLE OF INVENTION: NISIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP

STREET: 2101 L Street N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/773,731A  
FILING DATE: 24-DEC-1996  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/313,123  
FILING DATE: 18-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/00676  
FILING DATE: 01-APR-1993  
APPLICATION DATA:  
APPLICATION NUMBER: GB 9207267.7  
FILING DATE: 02-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Brady, Jr., James W.  
REGISTRATION NUMBER: 32,115  
REFERENCE/DOCKET NUMBER: E8280.016/P016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-785-9700  
TELEFAX: 202-887-0689  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-773-731A-2

Query Match 93.8%; Score 285; DB 3; Length 83;  
Best Local Similarity 94.7%; Pred. No. 8.2e-28;  
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57  
Db 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 14  
US-08-836-687B-41  
; Sequence 41, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-41

Query Match 92.9%; Score 282.5; DB 4; Length 56;  
Best Local Similarity 96.5%; Pred. No. 1.1e-27;  
Matches 55; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57  
Db 1 MSTKDNLDLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 56

RESULT 15

US-08-836-687B-47  
 ; Sequence 47 Application US/08836687B  
 ; Patent No. 6448034  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gasson, Michael John  
 ; APPLICANT: Dodd, Helen Malt  
 ; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
 ; FILE REFERENCE: 20747/70  
 ; CURRENT APPLICATION NUMBER: US/08/836,687B  
 ; CURRENT FILING DATE: 1995-11-20  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: Patent Ver. 2.1  
 ; SEQ ID NO 47  
 ; LENGTH: 57  
 ; TYPE: PRT  
 ; ORGANISM: Lactococcus sp.  
 -08-836-687B-47

Query Match 92.1%; Score 280; DB 4; Length 57;  
 Best Local Similarity 94.7%; Pred. No. 2.2e-27;  
 Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 MSTKDFNLDLVSVSKKSGASPRITSLCTPGCKTGALMGCMKMTATCCHSIHYSK 57  
 DB 1 MSTKDFNLDLVSVSKKSGASPRITSLCTPGCKTGALMGCMKMTATCCHSIHYSK 57

Search completed: June 7, 2003, 15:19:34  
 Job time : 15 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:14:21 ; Search time 20 Seconds  
(without alignments)  
294.235 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304  
Sequence: 1 MSTRDENLDVSVSKDSGA.....ALMGCMKMTATCHCSIHYSK 57

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 392085 seqs, 103240269 residues

tal number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PTI\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	60.9	34	10	US-09-030-619-206
2	185	60.9	34	10	US-09-017-340-14
3	157	51.6	56	10	US-09-030-619-208
4	73	24.0	708	9	US-10-184-644-211
5	73	24.0	708	9	US-10-184-634-211
6	70	23.0	4185	9	US-10-123-155-67
7	68.5	22.5	1570	9	US-10-184-644-335
8	68.5	22.5	1570	9	US-10-184-634-335
9	68.5	22.5	1570	9	US-10-184-634-335
10	68.5	22.5	1570	9	US-10-184-634-335
11	68.5	22.5	1570	9	US-10-184-634-335
12	68.5	22.5	1570	9	US-10-184-634-335
13	68.5	22.5	1570	9	US-10-184-634-335
14	68.5	22.5	1570	9	US-10-184-634-335
15	68.5	22.5	1570	9	US-10-184-634-335
16	68.5	22.5	1570	9	US-10-184-634-335
17	68.5	22.5	1570	9	US-10-184-634-335
18	68.5	22.5	1570	9	US-10-184-634-335
19	68.5	22.5	1570	9	US-10-184-634-335

20	68	22.4	4440	9	US-10-173-706-525	Sequence 525, App
21	68	22.4	4440	9	US-10-175-738-525	Sequence 525, App
22	68	22.4	4440	9	US-10-175-732-525	Sequence 525, App
23	68	22.4	4440	9	US-10-176-482-525	Sequence 525, App
24	68	22.4	4440	9	US-10-176-757-525	Sequence 525, App
25	68	22.4	4440	9	US-10-176-913-525	Sequence 525, App
26	68	22.4	4440	9	US-10-180-552-525	Sequence 525, App
27	68	22.4	4440	9	US-10-180-552-525	Sequence 525, App
28	68	22.4	4440	9	US-10-173-700-525	Sequence 525, App
29	68	22.4	4440	9	US-10-174-572-525	Sequence 525, App
30	68	22.4	4440	9	US-10-174-579-525	Sequence 525, App
31	68	22.4	4440	9	US-10-174-582-525	Sequence 525, App
32	68	22.4	4440	9	US-10-174-582-525	Sequence 525, App
33	68	22.4	4440	9	US-10-175-739-525	Sequence 525, App
34	68	22.4	4440	9	US-10-175-740-525	Sequence 525, App
35	68	22.4	4440	9	US-10-176-743-525	Sequence 525, App
36	68	22.4	4440	9	US-10-176-488-525	Sequence 525, App
37	68	22.4	4440	9	US-10-176-492-525	Sequence 525, App
38	68	22.4	4440	9	US-10-176-747-525	Sequence 525, App
39	68	22.4	4440	9	US-10-176-750-525	Sequence 525, App
40	68	22.4	4440	9	US-10-176-985-525	Sequence 525, App
41	68	22.4	4440	9	US-10-176-987-525	Sequence 525, App
42	68	22.4	4440	9	US-10-176-991-525	Sequence 525, App
43	68	22.4	4440	9	US-10-176-992-525	Sequence 525, App
44	68	22.4	4440	9	US-10-176-993-525	Sequence 525, App
45	68	22.4	4440	9	US-10-184-658-525	Sequence 525, App

## ALIGNMENTS

```
RESULT 1
US-09-030-619-206
Sequence 206, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 206
LENGTH: 34
TYPE: PRT
ORGANISM: Lactococcus lactis
US-09-030-619-206
Query Match
Best Local Similarity 97.1% Score 185: DB 10: Length 34:
Matches 33: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
OY 24 ITSTSLCTPGCKRGALMGCMKMTATCHCSIHYSK 57
Db 1 ITSTSLCTPGCKRGALMGCMKMTATCHCSIHYSK 34
RESULT 2
US-09-017-340-14
Sequence 14, Application US/09017340
Patent No. US20020090369A1
GENERAL INFORMATION:
APPLICANT: Murphy, Christopher J.
APPLICANT: McNulty, Jonathan F.
APPLICANT: Reid, Ted W.
```

;; TITLE OF INVENTION: Transplant Media  
;; FILE REFERENCE: TPLANT-06468  
;; CURRENT APPLICATION NUMBER: US/09/917,340  
;; CURRENT FILING DATE: 2001-07-29  
;; PRIOR APPLICATION NUMBER: 60/221,632  
;; PRIOR FILING DATE: 2000-07-28  
;; PRIOR APPLICATION NUMBER: 60/249,602  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/290,932  
;; PRIOR FILING DATE: 2001-05-15  
;; NUMBER OF SEQ ID NOS: 96  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 14  
;; LENGTH: 34  
;; TYPE: PRT  
;; ORGANISM: Lactococcus lactis  
US-09-917-340-14

Query Match 60.9%; Score 185; DB 10; Length 34;  
Best Local Similarity 97.1%; Pred. No. 5.5e-15;  
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 1 ITSISLCTPGCKTGALMGCMNMTATCHCSIHVS 34

RESULT 3  
US-09-030-619-208  
; Sequence 208, Application US/09030619B  
; Patent No. US20020035061A1  
; GENERAL INFORMATION:  
; APPLICANT: Krieger, Timothy J.  
; APPLICANT: Taylor, Robert  
; APPLICANT: Erfile, Douglas  
; APPLICANT: Fraser, Janet R.  
; APPLICANT: West, Michael H.P.  
; APPLICANT: Monicoll, Patricia J.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION  
; TITLE OF INVENTION: WITH ANTIBIOTICS  
; FILE REFERENCE: 660081.406  
; CURRENT APPLICATION NUMBER: US/09/030,619B  
; CURRENT FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 208  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-09-030-619-208

Query Match 51.6%; Score 157; DB 10; Length 56;  
Best Local Similarity 59.2%; Pred. No. 1.8e-11;  
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
Db 6 DFDLDVYKVS KODSKITTPQMKSESLCTPGCVTGALDTCFLQTLTCKCKI 54

RESULT 4  
US-10-184-644-211  
; Sequence 211, Application US/10184644  
; Publication No. US20030044930A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jilan  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James

;; APPLICANT: Smith, Victoria  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3430R1C227  
;; CURRENT APPLICATION NUMBER: US/10/184,644  
;; CURRENT FILING DATE: 2002-06-28  
;; PRIOR APPLICATION removed - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612  
;; SEQ ID NO: 211  
;; LENGTH: 708  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-184-644-211

Query Match 24.0%; Score 73; DB 9; Length 708;  
Best Local Similarity 43.2%; Pred. No. 1.9;  
Matches 16; Conservative 4; Mismatches 15; Indels 2; Gaps 1;  
Db 341 TGACCGCTTCTCTGCGATGACCGCACCTGTGGCGCT 377

RESULT 5  
US-10-184-634-211  
; Sequence 211, Application US/10184634  
; Publication No. US20030068684A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jilan  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C217  
; CURRENT APPLICATION NUMBER: US/10/184,634  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO: 211  
; LENGTH: 708  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-184-634-211

Query Match 24.0%; Score 73; DB 9; Length 708;  
Best Local Similarity 43.2%; Pred. No. 1.9;  
Matches 16; Conservative 4; Mismatches 15; Indels 2; Gaps 1;  
Db 341 TGACCGCTTCTCTGCGATGACCGCACCTGTGGCGCT 377

RESULT 6  
US-10-123-155-67  
; Sequence 67, Application US/10123155  
; Publication No. US20030068794A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc

```
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQUENCE ID NO 335
LENGTH: 1570
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-634-335
```

```
Query Match 23.5%; Score 68.5; DB 9; Length 1570;
Best Local Similarity 46.4%; Pred. No. 31;
Matches 13; Conservative 2; Mismatches 11; Indels 2; Gaps 1;
```

```
OY 25 TSTSLCTPGCKTGALMGCKNKTATCCHC 52
DB 33 TGTATCTTCTGCTGG--GCTATCTTCCCT 58
```

```
RESULT 7
US-10-184-644-335
Sequence 335, Application US/10184644
Publication No. US2003004930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQUENCE ID NO 335
LENGTH: 1570
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-335
```

```
Query Match 22.5%; Score 68.5; DB 9; Length 1570;
Best Local Similarity 41.2%; Pred. No. 16;
Matches 14; Conservative 2; Mismatches 15; Indels 3; Gaps 1;
```

```
OY 18 SGASPRITSTSLCTPGCKTGALMGCKNKTATCCHC 51
DB 918 TGAGCTCAATGCTGCGAAGATGCG---AGTCAC 948
```

```
RESULT 8
US-10-184-634-335
Sequence 335, Application US/10184634
Publication No. US20030068684A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQUENCE ID NO 335
LENGTH: 1570
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-634-335
```

```
Query Match 22.5%; Score 68.5; DB 9; Length 1570;
Best Local Similarity 41.2%; Pred. No. 16;
Matches 14; Conservative 2; Mismatches 15; Indels 3; Gaps 1;
```

```
OY 18 SGASPRITSTSLCTPGCKTGALMGCKNKTATCCHC 51
DB 918 TGAGCTCAATGCTGCGAAGATGCG---AGTCAC 948
```

```
RESULT 9
US-10-184-644-235
Sequence 235, Application US/10184644
Publication No. US2003004930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQUENCE ID NO 235
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-235
```

```
Query Match 22.5%; Score 68.5; DB 9; Length 1617;
Best Local Similarity 45.7%; Pred. No. 16;
Matches 16; Conservative 3; Mismatches 11; Indels 5; Gaps 2;
```

```
OY 18 SGASPRITSTSLCTPGCKTGALMGCKNKTATCCHC 52
DB 918 TGAGCTCAATGCTGCGAAGATGCG---AGTCAC 948
```

Db 971 AGAACAATTCCTCCTGCTTGA-TGC-----AACACT 1000

## RESULT 10

US-10-184-634-235  
Sequence 235, Application US/10184634  
Publication No. US20030068684A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C217  
CURRENT APPLICATION NUMBER: US/10/184,634  
CURRENT FILING DATE: 2002-06-28  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 235  
LENGTH: 1617  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-184-634-235

Query Match 22.5%; Score 68.5; DB 9; Length 1617;  
Best Local Similarity 45.7%; Pred. No. 16;  
Matches 16; Conservative 3; Mismatches 11; Indels 5; Gaps 2;

Qy 18 SGASPRITSTSLCTPGCKTGALMGCMKMTATC 52

Db 971 AGAACAATTCCTCCTGCTTGA-TGC-----AACACT 1000

RESULT 11  
US-10-123-155-317  
Sequence 317, Application US/10123155  
Publication No. US20030068794A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C30  
CURRENT APPLICATION NUMBER: US/10/123,155  
CURRENT FILING DATE: 2002-04-15  
Prior Application removed - See File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 317  
LENGTH: 1675  
TYPE: DNA

ORGANISM: Homo Sapien  
US-10-123-155-317

Query Match 22.5%; Score 68.5; DB 9; Length 1675;  
Best Local Similarity 48.4%; Pred. No. 19;  
Matches 15; Conservative 2; Mismatches 11; Indels 3; Gaps 1;

Qy 19 GASPRITSTSLCTPGCKTGALMGCMKMTATC 49

Db 476 GACTGCCATCACTTGCTG---GCTCCATTC 503

## RESULT 12

US-10-184-644-147  
Sequence 147, Application US/10184644  
Publication No. US20030044930A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C227  
CURRENT APPLICATION NUMBER: US/10/184,644  
CURRENT FILING DATE: 2002-06-28  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 147  
LENGTH: 1660  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-184-644-147

Query Match 22.4%; Score 68; DB 9; Length 1660;  
Best Local Similarity 37.5%; Pred. No. 19;  
Matches 12; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

Qy 18 SGASPRITSTSLCTPGCKTGALMGCMKMTATC 49

Db 735 AGTGATTCATCTTGCTGCCAGCAATGAC 766

## RESULT 13

US-10-184-634-147  
Sequence 147, Application US/10184634  
Publication No. US20030068684A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C217  
CURRENT APPLICATION NUMBER: US/10/184,634  
CURRENT FILING DATE: 2002-06-28  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612



SEQ ID NO 147  
LENGTH: 1660  
TYPE: DNA  
ORGANISM: Homo Saplen  
US-10-184-634-147

Query Match 22.4% Score 68; DB 9; Length 1660;  
Best Local Similarity 37.5% Pred. No. 19;  
Matches 12; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

OY 18 SGASPRITSLCTPGCKTGALMGNMKTATC 49  
DB 735 ACTGATTCATCTCTGCGTGCACCAATGAC 766

RESULT 14  
US-10-123-155-419  
Sequence 419, Application US/10123155  
Publication No. US20030068794A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Goddard, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumes, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C30  
CURRENT APPLICATION NUMBER: US/10/123, 155  
CURRENT FILING DATE: 2002-04-15  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 419  
LENGTH: 1781  
TYPE: DNA  
ORGANISM: Homo Saplen  
US-10-123-155-419

Query Match 22.4% Score 68; DB 9; Length 1781;  
Best Local Similarity 38.2% Pred. No. 21;  
Matches 13; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

OY 19 GASPRITSLCTPGCKTGALMGNMKTATCHCS 52  
DB 163 GCATCATGACCTCGCGGACTCTGAGAGTCTCA 196

RESULT 15  
US-10-184-644-233  
Sequence 233, Application US/10184644  
Publication No. US20030044930A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Goddard, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C227  
CURRENT APPLICATION NUMBER: US/10/184, 644  
CURRENT FILING DATE: 2002-06-28  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 233  
LENGTH: 2162  
TYPE: DNA  
ORGANISM: Homo Saplen  
US-10-184-644-233

Query Match 22.4% Score 68; DB 9; Length 2162;  
Best Local Similarity 44.1% Pred. No. 26;  
Matches 15; Conservative 2; Mismatches 15; Indels 2; Gaps 1;

OY 18 SGASPRITSLCTPGCKTG--ALMGNMKTATC 49  
DB 942 TGCCATCTATGACTTGCCTGACACCGCATGATC 975

Search completed: June 7, 2003, 15:18:51  
Job time : 21 secs

